

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 10:48:06 ; Search time 1871.78 seconds
(without alignments) 10979.953 Million cell updates/sec

Title: US-09-674-379A-15
Perfect-score: 1269
Sequence: 1 cagtcagcagtcagtcgttga.....atgtgtcagtcagtcacatc 1269

Scoring table: IDENTITY_NUC
Gap 10.0, Gape 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
EST:*
1: em_estb1:*
2: em_estb2:*
3: em_estb3:*
4: em_estb4:*
5: em_estb5:*
6: em_estb6:*
7: em_estb7:*
8: em_estb8:*
9: em_estb9:*
10: em_estb10:*
11: em_estb11:*
12: em_estb12:*
13: em_estb13:*
14: em_estb14:*
15: em_estb15:*
16: em_estb16:*
17: em_estb17:*
18: em_estb18:*
19: em_estb19:*
20: em_estb20:*
21: em_estb21:*
22: em_estb22:*
23: em_estb23:*
24: em_estb24:*
25: em_estb25:*
26: em_estb26:*
27: em_estb27:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	803.6	63.3	996	BQ720516	BQ720516 AGENCOURT
2	792.2	62.4	904	BQ716569	BQ716569 AGENCOURT
3	791	62.3	1014	BM921371	BM921371 AGENCOURT
4	771	60.8	954	BM545444	BM545444 AGENCOURT
5	666.8	52.5	891	BF339448	BF339448 AGENCOURT
6	665	52.4	822	BG109391	BG109391 602280448

7	639.2	50.4	796	13	BI453348	BI453348 603173067
8	636	50.1	644	13	BC938621	BC938621 CN27C11.X
9	592.8	46.7	717	13	BI251219	BI251219 602994427
10	576.4	45.4	933	14	BQ716088	BQ716088 AGENCOURT
11	573.6	45.2	627	12	BF090699	BF090699 QV2-NT004
12	566.2	44.6	882	13	BI825121	BI825121 603038662
13	565.2	44.5	870	14	BQ930927	BQ930927 AGENCOURT
14	564.4	44.5	912	13	BI768516	BI768516 603057118
15	549.6	43.3	880	13	BI764020	BI764020 603043221
16	544	42.9	835	14	BQ718885	BQ718885 AGENCOURT
17	540.2	42.6	627	13	BI183862	BI183862 UNL-P-FN-
18	537.8	42.4	745	12	BC431026	BC431026 602498636
19	510	40.2	531	10	AW957458	AW957458 EST369423
20	507.4	40.0	794	13	BI837271	BI837271 603090563
21	495.8	39.1	815	12	BI174360	BI174360 602336371
22	486	38.3	487	9	AL041285	AL041285 DKFZP34M
23	480	37.8	721	13	BI523975	BI523975 603052129
24	479	37.7	491	14	BM707581	BM707581 UI-E-C11-
25	477.8	37.7	494	14	BO695207	BO695207 1001350.H
26	459.4	36.2	1006	9	AL545549	AL545549 AL545549
27	446.2	35.2	762	13	BI834712	BI834712 603090304
28	442	34.8	690	12	BF104873	BF104873 601822595
29	415.4	32.7	576	10	AW142226	AW142226 EST292462
30	405.4	31.9	889	14	BO897144	BO897144 AGENCOURT
31	387.8	30.6	391	14	BM699401	BM699401 UI-E-DX1-
32	384.2	30.3	462	10	BE656312	BE656312 UI-M-BH0-
33	381.2	30.0	685	9	AL543659	AL543659 AL543659
34	377	29.7	629	13	BI489863	BI489863 603031610
35	375.4	29.6	377	14	N95751	N95751 zbf5e06.s1
36	374.2	29.5	1015	13	BM544363	BM544363 AGENCOURT
37	371.2	29.3	1850	11	BC018871	BC018871 Homo sapi
38	370.4	29.2	389	10	BE004365	BE004365 CMO-BN010
39	366.6	28.9	430	10	AW744572	AW744572 ur32h12.Y
40	359.2	28.3	906	14	BQ723059	BQ723059 AGENCOURT
41	358.2	28.2	443	9	AA801465	AA801465 EST190962
42	356	28.1	477	9	AA037185	AA037185 zc51c05.x
43	355.4	28.0	470	10	AW572176	AW572176 xt75e10.X
44	353	27.8	586	14	BM695168	BM695168 UI-E-CQ1-
45	349	27.5	371	12	BF090673	BF090673 QV2-NT004

ALIGNMENTS

RESULT 1
LOCUS BQ720516 996 bp mRNA linear EST 16-JUL-2002
DEFINITION AGENCOURT.8219503 lupsk1.sympathetic_trunk Homo sapiens cDNA clone
IMAGE:6186164 5', mRNA sequence.
ACCESSION BQ720516
VERSION BQ720516.1 GI:21859413
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 996)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNML Unpublished (1999)
COMMENT Contact: Robert Strussberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM13583 row: f column: 05
High quality sequence stop: 634.
Location/Qualifiers
1..996

FEATURES
SOURCE

BASE COUNT	ORIGIN
245 a	284 c
242 g	224 t
1	others

Query Match	Best Local	Similarity	Score	DB	Length
Matches	866	Conservative	0	Mis-Matches	10
				Indels	Gaps
QY	1	CAGTGCACGATGCGCTTTGACCTGGATGCGCCGATGCGACAGTGTGTTAACCATAATGGCGG	60		
DB	26	CAGTGCACGATGCGCTTTGACCTGGATGCGCCGATGCGACAGTGTGTTAACCATAATGGCGG	120		
QY	61	TGCCCAACCATCCCGGAGCGCTGCCGAGAGACATGATGTGTGTTAACCATAATGGCGG	145		
DB	86	TGCCCAACCATCCCGGAGCGCTGCCGAGAGACATGATGTGTGTTAACCATAATGGCGG	145		
QY	121	TATTTATGCAATCCCGGAGCAAAACCGTGTATGAGAGGCGCTTACTGAACCCCTACTCG	180		
DB	146	TATTTATGCAATCCCGGAGCAAAACCGTGTATGAGAGGCGCTTACTGAACCCCTACTCG	205		
QY	181	ACCCCTTACTAGTCCGTACCCAGACAGTCCGCCACACTCTAGCTCCAAACTATGCC	240		
DB	206	ACCCCTTACTAGTCCGTACCCAGACAGTCCGCCACACTCTAGCTCCAAACTATGCC	265		
QY	241	ACGATCTCCAGCGCTTATATATGCGCGCTTTGGATACCAATGATGATGATAAGCAACCAATGT	300		
DB	266	ACGATCTCCAGCGCTTATATATGCGCGCTTTGGATACCAATGATGATGATAAGCAACCAATGT	325		
QY	301	GTGATGTGACGAGTGTGCMAACGATTTCCACAGTGCACACCCACCCGAGTCTGCATC	360		
DB	326	GTGATGTGACGAGTGTGCMAACGATTTCCACAGTGCACACCCACCCGAGTCTGCATC	385		
QY	361	AATCTGAAGGCGGCTACACCTGCTCTGCAACGAGATATTTGGCTTGTGAAAGCCAG	420		
DB	386	AATCTGAAGGCGGCTACACCTGCTCTGCAACGAGATATTTGGCTTGTGAAAGCCAG	445		
QY	421	TGCTTAGACATTTGATGTGCTATATGTTTACTGACAGAGTCTGTGCGAATGTTCT	480		
DB	446	TGCTTAGACATTTGATGTGCTATATGTTTACTGACAGAGTCTGTGCGAATGTTCT	505		
QY	481	GGATCCATCTTGTACATGCAACCCCTGTTTTACCTCAATGAGATGGAAGTCTTGC	540		
DB	506	GGATCCATCTTGTACATGCAACCCCTGTTTTACCTCAATGAGATGGAAGTCTTGC	565		
QY	541	CAAGATGTGAACGAGTGTGCACACCGGAACCCCTGCGTCAAACTGGCTCAACCTTAC	600		
DB	566	CAAGATGTGAACGAGTGTGCACACCGGAACCCCTGCGTCAAACTGGCTCAACCTTAC	625		
QY	601	GGCTTTTCACTGCGCGTGTACCCAGAGATATGAATTGAGAGATGGGTTCAATGCG	660		
DB	626	GGCTTTTCACTGCGCGTGTACCCAGAGATATGAATTGAGAGATGGGTTCAATGCG	685		
QY	661	AGGATATGACGAGTGCAGCTTCTCTGAGTCTCTGCCAACAATGAGTGTGTGACACG	720		
DB	686	AGGATATGACGAGTGCAGCTTCTCTGAGTCTCTGCCAACAATGAGTGTGTGACACG	745		

QY	721	CCGGGCATATTTTGGCTCCGCC-TCGAGCTCATCTCGTGGATGACACCGAAG	779
Db	746	CCGGGCATATTTTGGCTCCGCCCTTCAGCTCATCTCGTGGATGACACCGAAG	805
QY	780	CTGGCAACATCAACGAATGTGACACAGG-AACACACGT-GCAACTGCAGAG-AC	836
Db	806	CTGCCAACATCAACGAATGTGACACAGGAACACACGTGGCAACTGCAGAGAAC	865
QY	837	GTGCTACATTT--ACAAGGGGGCTTCAAATGCATGCAGCCC	876
Db	866	GTGCTACAAATTTACCAAGGGGGCTTTAAATGCATGCAGACC	908

RESULT 2		
B0716569		
LOCUS	B0716569	904 bp
DEFINITION	AGENOUR-8215248 lupsk1-sympathetic-trunk Homo sapiens EST 16-JUL-2002	mRNA linear
IMAGE:	6187340 5'	MRNA sequence.
ACCESSION	B0716569	
VERSION	B0716569.1	GI:2185466
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	

REFERENCE	Mammals: Eutheria; Primates; Carnivata; Vertebrata; Euteleostomi;
AUTHORS	I (bases 1 to 904)
TITLE	NIH-MGC http://mgc.nhl.nih.gov/ .
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1999) Contact: Robert Strausberg Ph.D.

ISSUE Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MCC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLNL13581 row: c column: 21
High quality sequence stop: 646.

FEATURES	Source
Location/Qualifiers	1. .904
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/clone_image:6187340"	
/clone_lib="lupski_sympathetic_trunk"	
/sex="male"	
/tissue_type="sympathetic trunk"	
/dev_stage="adult, 16 yr"	
/lab_host="DH10B"	
/note="Vector: pCMV-SPORT6 (Life Technologies); Site:1: Not1; Site:2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCGACGCGCTCG-3' and 5'-GACGACGTCCTGATGCGCGACGCGCCGCT(15)-3'. Size selected > 1 kb for average insert length 1.9 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine); available through Life Technologies."	
	219 a 264 c 219 g 202 t

Query Match	62.4%	Score 792.2	DB 14	Length 904
Best Local Similarity	97.1%	Pred. No. 2.2e+205		
Matches	861	Conservative	0	Mismatches 18; Indels 8; Gaps 5
QY	13	GGCTTTGACCTGATGCGCAGTCAAGACAGGTTTAGATTATGATGATGGCAACATC	72	
Db	1	GGCTTTGACCTGATGCGCAGTCAAGACAGGTTTAGATTATGATGATGGCAACATC	60	
QY	73	CCCGAGGCCCTCCGAGAGACATGATGTGTGTTAACCAAAATGGCGGGTATTTATGCATT	132	

```

Db      61 CCCGAGCCCTGCGAGAGACATGATGTGTTAAACCAAAATGGCGGATTTATGAT 120
Qy      133 CCCCCGAAACCCCTGTATGTAGAGGCCCCCTACTGCAACCCCTACTGACCCCTACTCA 192
Db      121 CCCCCGAAACCCCTGTATGTAGAGGCCCCCTACTGCAACCCCTACTGACCCCTACTCA 180
Qy      193 GGTCCGTCACGACGAGTGGCCCCACACTCTCAAGTCCCAACATCCACGATCCAGG 252
Db      181 GGTCCGTCACGACGAGTGGCCCCACACTCTCAAGTCCCAACATCCACGATCCAGG 240
Qy      253 CCTCTTATATGCGCGTTTGATACAGATGATGAAGAACCAATGTGTGATGTGAC 312
Db      241 CCTCTTATATGCGCGTTTGATACAGATGATGAAGAACCAATGTGTGATGTGAC 300
Qy      313 GAGTGTGCAACAGATTTCCACAGAGTGCACACCCCAATCTGCATCAATACTGAAGG 372
Db      301 GAGTGTGCAACAGATTTCCACAGAGTGCACACCCCAATCTGCATCAATACTGAAGG 360
Qy      373 GGTGACACCTCTCTCTGACACGAGATATGAGCTCTGGAAGGCGAGTCTTAGACATT 432
Db      361 GGTGACACCTCTCTCTGACACGAGATATGAGCTCTGGAAGGCGAGTCTTAGACATT 420
Qy      433 GATGAATGTGCTATGATGTACTGACAGAGCTCTGTGCAATGTTCTGATCTTATCT 492
Db      421 GATGAATGTGCTATGATGTACTGACAGAGCTCTGTGCAATGTTCTGATCTTATCT 480
Qy      493 TGTACATGCAACCCGTTTACCCCTCATGTAGATGAAGTCTTGCCAAAGATGTGAC 552
Db      481 TGTACATGCAACCCGTTTACCCCTCATGTAGATGAAGTCTTGCCAAAGATGTGAC 540
Qy      553 GAGTGTGCAACGAGAACCCCTGCTGCAACCTGCTGAACACCTAGAGCTCTTTCATC 612
Db      541 GAGTGTGCAACGAGAACCCCTGCTGCAACCTGCTGAACACCTAGAGCTCTTTCATC 600
Qy      613 TGCCGCTGTGACCCAGATATGATGAGAAAGTGGCTTCATGCAATGATATGAC 672
Db      601 TGCCGCTGTGACCCAGATATGATGAGAAAGTGGCTTCATGCAATGATATGAC 660
Qy      673 GAGTGCAGCTCTCTGATGCTCTGCAACATGATGTGTAACGACCCGCGACATAC 732
Db      661 GAGTGCAGCTCTCTGATGCTCTGCAACATGATGTGTAACGACCCGCGACATAC 720
Qy      733 TTTGCTCTCTGCTCTGACAGCTACAT-CTGCTGATGACCAAC-CAAGCTGCAAGA-- 788
Db      721 TTTGCTCTCTGCTCTGACAGCTACAT-CTGCTGATGACCAAC-CAAGCTGCAAGA-- 780
Qy      789 CATCAAGCAATGTGACACAGGACCAACAGTGCAC-CTGCAAGACAGCTGTACATT 847
Db      781 ATCAACCGAAATGTGACACAGGACCAACAGTGCAC-CTGCAAGACAGCTGTACATT 840
Qy      848 T---ACAAGGGGGGCTTCAATGATGACCCCATCCGCTGTGAGAG 891
Db      841 TTACAAGGGGGGCTTCAATGATGACCCCATCCGCTGTGAGAG 887

```

RESULT 3
 BM921371
 LOCUS 1014 bp mRNA linear EST 12-MAR-2002
 DEFINITION AGENCOURT.6626255 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5752785
 5', mRNA sequence.
 ACCESSION BM921371
 VERSION BM921371.1 GI:19371750
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 1014)
 NIH-MGC <http://mgi.nci.nih.gov/>;
 AUTORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing By: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
<http://image.llnl.gov>
 Plate: L14M12787 row: 1 column: 10
 High quality sequence stop: 738.
 Location/Qualifiers
 1..1014
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5752785"
 /clone_1ib="NIH_MGC_115"
 /lab_host="DH10B"
 /note="Organ: pooled brain, lung, testis; Vector:
 pCMV-SPORT6; Site:1: NotI; Site:2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.8 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 021. Note: this is a NIH MGC Library."

```

BASE COUNT 235 a 300 c 246 g 233 t
ORIGIN
Query Match 62.3%; Score 791; DB 14; Length 1014;
Best Local Similarity 98.2%; Pred. No. 4.9e-205;
Matches 842; Conservative 0; Mismatches 10; Indels 5; Gaps 4;
Qy      1 CAGTGCACGAATGGCTTTGACCTGATCCGACGACGACAGTGTATGATTTGATGATA 60
Db      163 CAGTGCACGAATGGCTTTGACCTGATCCGACGACGACAGTGTATGATTTGATGATA 222
Qy      61 TGCCGCAACATCCCGAGGCGCTGCGAGAGACATGATGTGTTAACCAAAATGGCGGG 120
Db      223 TGCCGCAACATCCCGAGGCGCTGCGAGAGACATGATGTGTTAACCAAAATGGCGGG 282
Qy      121 TATTTATGATTTCCCGGACAAACCTGTGTATGAGAGGCGCTACTGCAACCCCTACTCG 180
Db      283 TATTTATGATTTCCCGGACAAACCTGTGTATGAGAGGCGCTACTGCAACCCCTACTCG 342
Qy      181 ACCCCCTACTCAGATCCGTTACCCAGCAGAGCTGCCACACCTCTAGCTCCAACTATCCC 240
Db      343 ACCCCCTACTCAGATCCGTTACCCAGCAGAGCTGCCACACCTCTAGCTCCAACTATCCC 402
Qy      241 ACGATCTCCAGGCGCTTTATATGCGGCTTTGATACAGATGATGAAGCAACCAATGT 300
Db      403 ACGATCTCCAGGCGCTTTATATGCGGCTTTGATACAGATGATGAAGCAACCAATGT 462
Qy      301 GTGATGTGAGACAGTGTGCAACAGATTTCCACAGAGTGCACACCCACCCAGATCTGATC 360
Db      463 GTGATGTGAGACAGTGTGCAACAGATTTCCACAGAGTGCACACCCACCCAGATCTGATC 522
Qy      361 AATACTGAAGGGGGGTACACCTCTCTGACACGAGGATATGTTGCTTCTGGAAGGCCAG 420
Db      523 AATACTGAAGGGGGGTACACCTCTCTGACACGAGGATATGTTGCTTCTGGAAGGCCAG 582
Qy      421 TGCTTAGACATGATGAATGTGCTATGTTACTGACGAGAGCTGTGCGCAATGTTCT 480
Db      583 TGCTTAGACATGATGAATGTGCTATGTTACTGACGAGAGCTGTGCGCAATGTTCT 642
Qy      481 GGATCCTATCTTGTACATGCAACCCCTGTTTAACTCTCAATGAGATGAGAGTCTTGC 540
Db      643 GGATCCTATCTTGTACATGCAACCCCTGTTTAACTCTCAATGAGATGAGAGTCTTGC 702
Qy      541 CAAGATGTAAAGAGTGTGCAACGAGAACCCCTGCTGCAAACTGCTGAACACCTAC 600
Db      703 CAAGATGTAAAGAGTGTGCAACGAGAACCCCTGCTGCAAACTGCTGAACACCTAC 762

```

QY 601 GGCTCTTTCATCTCCCGCTGTGACCCAGGATATGAACCTTGAGAGAGATGGCTTCATTCG 660
 DB 763 GGCTCTTTCATCTCCCGCTGTGACCCAGGATATGAACCTTGAGAGAGATGGCTTCATTCG 822
 QY 661 AGTGAATATGACGAGTGCAGCTTCTCTGATTCCTCTGCGCAACATGATGTGTGAACCG 720
 DB 823 AGTGAATATGACGAGTGCAGCTTCTCTGATTCCTCTGCGCAACATGATGTGTGAACCG 882
 QY 721 CCGGGCAGCATCTCTGCTCTGCGCTGCACTTCCTGATGTACACCGAAGC 780
 DB 883 CCGGGCAGCATCTCTGCTCTGCGCTGCACTTCCTGATGTACACCGAAGC 941
 QY 781 TGGCAAGACATCAACGATGTGACACAGAACACACAGTGCAGCTGAGAGAGAGAG 840
 DB 942 TGGCAAGACATCAACGATGTGACACAGAACACACAGTGCAGCTGAGAGAGAGAG 997
 QY 841 TACAAATTTACAGGGGG 857
 DB 998 TACAAATTTACAGGGGG 1014

RESULT 4 BM545444

DEFINITION AGENCOURT_6500472 NIH_MGC_124 Homo sapiens CDNA clone IMAGE:5729596
 5', mRNA sequence.

ACCESSION BM545444
 VERSION BM545444.1 GI:18777513
 KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.

Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: Invitrogen

CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/ILNL at:
 http://image.llnl.gov

Plate: LLAM12727 row: C column: 05
 High quality sequence stop: 676.

Location/Qualifiers
 1..954

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5729596"
 /clone_1b="NIH_MGC_124"
 /tissue_type="hippocampus"
 /lab_host="DH10B"

/note="Organ: brain; Vector: PCMV-SPORT6; Site: 1; Ecorev
 (destroyed); Site: 2; Note: RNA source male hippocampus,
 age 27. Library is oligo-dT primed and directionally
 cloned (Ecorev site is destroyed upon cloning). Average
 insert size 1.4 kb, insert size range 0.9-4 kb. Library is
 normalized and enriched for full-length clones and was
 constructed by C. Gruber (Invitrogen). Research Genetics
 tracking code 012."

BASE COUNT 232 a 278 c 238 g 206 t

ORIGIN

Query Match 60.8%; Score 771; DB 13; Length 954;
 Best Local Similarity 100.0%; Pred. No. 1.4e-199;
 Matches 771; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

499 TGCACCCCTGGTTTACCCATGAGAGTGAAGGCTTCCCAAGATGTACAGAGTGT 558
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 8 TGCACCCCTGGTTTACCCATGAGAGTGAAGGCTTCCCAAGATGTACAGAGTGT 67
 QY 559 GCCACCGAAGAACCCCTGGCTGCAACCTGCGTCAACACCTACAGGCTTTTCATCTGCGCG 618
 DB 68 GCCACCGAAGAACCCCTGGCTGCAACCTGCGTCAACACCTACAGGCTTTTCATCTGCGCG 127
 QY 619 TGTGACCGAGATATGAACTTGAGAGATGGCGCTTCAATGATGATGTGACAGAGTGC 678
 DB 128 TGTGACCGAGATATGAACTTGAGAGATGGCGCTTCAATGATGATGTGACAGAGTGC 187
 QY 679 AGCTTCTGTGAGTCTCTGCGCAACATGATGTGTGAACACCGCGGCACTATCTTCTGC 738
 DB 188 AGCTTCTGTGAGTCTCTGCGCAACATGATGTGTGAACACCGCGGCACTATCTTCTGC 247
 QY 739 TCTGCGCCCTCAGGCTATCTCTGCTGATGACAAACGAGTGCACACATCAACGAA 798
 DB 248 TCTGCGCCCTCAGGCTATCTCTGCTGATGACAAACGAGTGCACACATCAACGAA 307
 QY 799 TGTGACCGAAGAACACAGTGCAGACCTGAGAGAGCTTATCTGAGATGATGATTAACG 858
 DB 308 TGTGACCGAAGAACACAGTGCAGACCTGAGAGAGCTTATCTGAGATGATGATTAACG 367
 QY 859 TTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 918
 DB 368 TTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 427
 QY 919 TGTATGTCTCTGCTGAGAACCTTGGCTGCGAGACAGCCCTTTACATCTTGTACCGG 978
 DB 428 TGTATGTCTCTGCTGAGAACCTTGGCTGCGAGACAGCCCTTTACATCTTGTACCGG 487
 QY 979 GACATGAGAGTGTGATGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1038
 DB 488 GACATGAGAGTGTGATGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 547
 QY 1039 ACCCGCTACCTGCGGCTTATTCATTTCCAGATCAATCTGGGAATGAGAGGAGAGAA 1098
 DB 548 ACCCGCTACCTGCGGCTTATTCATTTCCAGATCAATCTGGGAATGAGAGGAGAGAA 607
 QY 1099 TTTTACATGCGGGAAGAGGCGCCCATCATGATCCACCTGCGTGTATGACAGCCCATCAAA 1158
 DB 608 TTTTACATGCGGGAAGAGGCGCCCATCATGATCCACCTGCGTGTATGACAGCCCATCAAA 667
 QY 1159 GGGCCCCGGGAATTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1218
 DB 668 GGGCCCCGGGAATTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 727
 QY 1219 AGAGGACGCTCGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1269
 DB 728 AGAGGACGCTCGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 778

RESULT 5 BF339448

LOCUS BF339448 891 bp mRNA linear EST 22-NOV-2000
 DEFINITION 602039102F1 NCI_CGAP_Brn64.Homo sapiens CDNA clone IMAGE:4186704
 5', mRNA sequence.

ACCESSION BF339448
 VERSION BF339448.1 GI:11285903
 KEYWORDS EST.

SOURCE human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.

Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: David N. Louis, M.D.

CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LHAM9508 row: d column: 01
 High quality sequence stop: 675.
 Location/Qualifiers

FEATURES
SOURCE

1. 891
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="NCI-GAP_Brn64"
 /tissue_type="g1oblastoma with EGFR amplification"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: brain; Vector: pCMV-SPO6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.57 kb. Constructed by Life Technologies. Note: this is a NCI-GAP Library."
 222 a 258 c 230 g 181 t

BASE COUNT
ORIGIN

Query Match 52.5%; Score 666.8; DB 12; Length 891;
 Best Local Similarity 95.1%; Pred. No. 4e-171;
 Matches 742; Conservative 0; Mismatches 32; Indels 6; Gaps 5;

229 CCAACTATCCACGATCTCCAGGCTCTTATATGCGCTTGGATACCATGATGAA 288
 1 CCAACTATCCACGATCTCCAGGCTCTTATATGCGGCTTGGATACCATGATGAA 59
 289 AGCAACCAATGTGTGATGTGAGAGAGTGTGACAGATTCACAGATGCAACCCACC 348
 60 AGCAACCAATGTGTGATGTGAGAGAGTGTGACAGATTCACAGATGCAACCCACC 119
 349 CAGATCTGCATCAATCACTGAAGGCGGTACACCTGCTCTGACCGAGATATGGCTT 408
 120 CAGATCTGCATCAATCACTGAAGGCGGTACACCTGCTCTGACCGAGATATGGCTT 178
 409 CTGGAAGGCGGCTGTAGACATGTGATGATGTGCTATGCTAGTCCAGACGCTGTG 468
 179 CTGGAAGGCGGCTGTAGACATGTGATGATGTGCTATGCTAGTCCAGACGCTGTG 238
 469 GCGAATGTCTGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTG 528
 239 GCGAATGTCTGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTG 298
 529 GGAAGGCTTGCAGATGTGAGAGAGTGTGACAGATTCACAGATGCAACCCACC 588
 299 GGAAGGCTTGCAGATGTGAGAGAGTGTGACAGATTCACAGATGCAACCCACC 358
 589 GTCAACACCTACGCTCTTTCATCTGCGCTGTGACCGAGATATGAATGAGAGAT 648
 359 GTCAACACCTACGCTCTTTCATCTGCGCTGTGACCGAGATATGAATGAGAGAT 418
 649 GCGCTTCAATGAGTGTGATGTGAGAGAGTGTGAGTGTGCTGTGCAACATGAG 708
 419 GCGCTTCAATGAGTGTGATGTGAGAGAGTGTGAGTGTGCTGTGCAACATGAG 478
 709 TGTGTGAACCGCCCGGAGATCTTCTGCTGCGCCCGAGGCTACATCTGCTGGAT 768
 479 TGTGTGAACCGCCCGGAGATCTTCTGCTGCGCCCGAGGCTACATCTGCTGGAT 538
 769 GACAACGAGCTGCGCAAGATCAACGATGTGAGACAGAGAACACAGCTCAACCTG 828
 539 GACAACGAGCTGCGCAAGATCAACGATGTGAGACAGAGAACACAGCTCAACCTG 598
 829 CAGCAGAGCTGCTCAATTTACAGGGGCTTCAAA-TGCATGACCCCATCCGCTGTA 887
 599 CAGCAGAGCTGCTCAATTTACAGGGGCTTCAAA-TGCATGACCCCATCCGCTGTA 658
 888 GGAGCTTATCTGAGAGATCAAGATACCGCT-GTATGTGCTGCTGTAACACCTGGCT 946
 659 GGAGCTTATCTGAGAGATCAAGATACCGCTGTAACACCTGGCTGTAACACCTGGCT 716
 947 GCAGAGACGAGCTTATCTGAGAGATCAAGATACCGCTGTAACACCTGGCTGTAACACCTGGCT 1006

DB

717 GCGAACAACAGCCCTTACCATCTGTCCGGACATGCGCCCTGGGAGAGAGCGGCGG 776

RESULT 6

LOCUS
 DEFINITION

BG109391 822 bp mRNA linear EST 30-JAN-2001
 602280448F1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4368227 5',
 mRNA sequence.

ACCESSION

VERSION
 KEYWORDS

SOURCE
 ORGANISM

REFERENCE
 AUTHORS

TITLE
 JOURNAL

COMMENT
 Email: cgapds-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LHAM10021 row: 0 column: 12
 High quality sequence stop: 676.
 Location/Qualifiers

FEATURES
SOURCE

1. 822
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="NIH_MGC_86"
 /tissue_type="osteosarcoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: bone; Vector: pCMV-SPO6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; Oligo-dt primed. Average insert size 1.533 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
 204 a 227 c 210 g 181 t

BASE COUNT
ORIGIN

Query Match 52.4%; Score 665; DB 12; Length 822;
 Best Local Similarity 97.8%; Pred. No. 1.2e-170;
 Matches 717; Conservative 0; Mismatches 10; Indels 6; Gaps 4;

493 TGTACATGCAACCGCTGTTTACCTTCATGAGATGAGAGGCTTGGCAAGATGTGAA 552
 5 TGTACATGCAACCGCTGTTTACCTTCATGAGATGAGAGGCTTGGCAAGATGTGAA 64
 553 GAGTGTGCAACGAGAACCCCTGCGCAACCGGCTGCAACACCTACGGCTTTCATC 612
 65 GAGTGTGCAACGAGAACCCCTGCGCAACCGGCTGCAACACCTACGGCTTTCATC 124
 613 TCCCGCTGTGACCGAGATATGAGATGAGAGATGAGAGGCTTTCATGAGATGTGAA 672
 125 TCCCGCTGTGACCGAGATATGAGATGAGAGATGAGAGGCTTTCATGAGATGTGAA 184
 673 GAGTGTGCAACCGCTGTTTACCTTCATGAGATGAGAGGCTTGGCAAGATGTGAA 732
 185 GAGTGTGCAACCGCTGTTTACCTTCATGAGATGAGAGGCTTGGCAAGATGTGAA 244
 733 TTTCTGCTCTGCGCCCGGCTACATCTGCTGTGATGAGAGGCTGCAAGATGTGAA 792
 245 TTTCTGCTCTGCGCCCGGCTACATCTGCTGTGATGAGAGGCTGCAAGATGTGAA 304
 793 AACGATGTGAGACAGAACCAACAGCTGAGAGGCTGAGAGGCTGCTACAAATTTACAA 852

BASE COUNT	195 a	208 c	197 g	196 t
ORIGIN				
Query Match	50.4%;	Score 639.2;	DB 13;	Length 796;
Best Local Similarity	89.2%;	Pred. No. 1.3e-163;		
Matches 712;	Conservative	0;	Mismatches 83;	Indels 3; Gaps 2;
Oy	279	GATGATGATGAAGCAACCAACCAATATGTGTGATGTGTGGACGAGTGTGTCACACAGATTTCCCAACCACTG	338	
Db	1	GATGATGATGAAGCAACCAACCACTGTGTGATGTGTGGACGAGTGTGTCACACAGACTACACCACTG	60	
Oy	339	CAACCCCAACCCAGATCTGCATCAATTAATTAAGGAGGCGGTACACACGCTCTCTCACGACGAG	398	
Db	61	CAACCCCTTACCCAGATCTGTATCAACACCTGAAAGGAGTTACACCTGTCTCTCACGACGATG	120	
Oy	399	ATATTGGCTTGTGAAAGGCGAGTCTTAGACATTGATGAATGTGCTATGTTACTGTGCA	458	
Db	121	GTACTGGCTTGTGAAAGGCGAGTGTGATGATTAATGATGAATGTGCTATGTTACTGTGCA	180	
Oy	459	GCACCTCTGTGCGAATGTCTCTGCATCTATTCTTGTACATGCAACCACTGTGTTTAACTT	518	
Db	181	GCACCTCTGTGCGAATGTCTCAGAGATCTATTCTGTACATGCAACCACTGTGTTTAACTT	240	
Oy	519	CAATGAGAGATGAAGGCTTGTGCCAAGATGTATACGAGTGTGCCACCGAGAACCCCTGCGT	578	
Db	241	CACGACAGATGAAGGCTTGTGCCAAGATGTATACGAGTGTGCCAAGTGAATTCCTGTGCT	300	
Oy	579	GCAAACTGTGTCACACACTTACGGGCTTTTCATCTGCCGTGTGACCCAGATATGAAT	638	

OY	639	TTGAGGAAGATGGGCGTTCATTTCGACGTGATATGAGAGAGAGGACGATTCTCTGAGTTCCCTTG	698
Db	361	TGAGGAAGATGGCATTCATCTCAGTGAATATGACAGAGTGCAGCTTCTCCGAGTTCCTGTG	420
OY	699	CCAACATGAGTGTGTGAACACGCCCGCACATCTTGCTGCTCGTCCCTCCAGGCTACAT	758
Db	421	TCAACAGCAGTGTGTGAACACGCCCGGCTCATCTTCTGCTCTGTCCTCCAGGTACGT	480
OY	759	CCTGCTGGATGACACACCGAGCTGCGCAAGACATCAACGATGTGAGCAGACGAACACAC	818
Db	481	CCTGTTGGATGATTAACCGAGCTCCAGAGATATCAATGAATGTGAGCACCGAAACACAC	540
OY	819	GTGCAACCTGAGAGAGACGCTCTCAATTTTCAAGGGGGGCTTCAATTCATCTGACGCCAT	878
Db	541	GTTGATCTCATCTGAGACTTGTCTTCAAACTTCAAGGGGGCTTCAATTAATGATTCAT	600
OY	879	CCGCTGTGAGAGCGCTTATCTGAGAGATCAGATTAACCGCTGTATGTCTCTGCGAG -A	937
Db	601	CAGCTGTGAGAGCGCTTATCTGCTGATTTGGTGAACCGCTGTATGTCTCTGCTGAGCA	660
OY	938	ACCCTGGCTGACAGACACGCGCTTTTACCATCTTGTACCGGGAGATGGACGTGTGTCAG	997
Db	661	ACACACAGTGCAGAGACACGCGCATTTACCATCTGTATCGGGACATGGATGTGGTCTAG	720
OY	998	GAGCGTCCGTTCCCGCTGACATCTTCCAAATGCAAGCACACACCCGCTACCTGGGGCT	1053
Db	721	GAGCGTCCGTTCCCGCTGACATCTTCCAGATGCAAGCAA -AACCAGTACTCTGTGGCT	778

Accession	Version	Keywords	Source	LOCUS	DEFINITION	Accession	Version	Keywords	Source
U00001	1	plasmid	human	U00001	U00001	U00001	1	plasmid	human
U00002	1	plasmid	human	U00002	U00002	U00002	1	plasmid	human
U00003	1	plasmid	human	U00003	U00003	U00003	1	plasmid	human
U00004	1	plasmid	human	U00004	U00004	U00004	1	plasmid	human
U00005	1	plasmid	human	U00005	U00005	U00005	1	plasmid	human
U00006	1	plasmid	human	U00006	U00006	U00006	1	plasmid	human
U00007	1	plasmid	human	U00007	U00007	U00007	1	plasmid	human
U00008	1	plasmid	human	U00008	U00008	U00008	1	plasmid	human
U00009	1	plasmid	human	U00009	U00009	U00009	1	plasmid	human
U00010	1	plasmid	human	U00010	U00010	U00010	1	plasmid	human
U00011	1	plasmid	human	U00011	U00011	U00011	1	plasmid	human
U00012	1	plasmid	human	U00012	U00012	U00012	1	plasmid	human
U00013	1	plasmid	human	U00013	U00013	U00013	1	plasmid	human
U00014	1	plasmid	human	U00014	U00014	U00014	1	plasmid	human
U00015	1	plasmid	human	U00015	U00015	U00015	1	plasmid	human
U00016	1	plasmid	human	U00016	U00016	U00016	1	plasmid	human
U00017	1	plasmid	human	U00017	U00017	U00017	1	plasmid	human
U00018	1	plasmid	human	U00018	U00018	U00018	1	plasmid	human
U00019	1	plasmid	human	U00019	U00019	U00019	1	plasmid	human
U00020	1	plasmid	human	U00020	U00020	U00020	1	plasmid	human
U00021	1	plasmid	human	U00021	U00021	U00021	1	plasmid	human
U00022	1	plasmid	human	U00022	U00022	U00022	1	plasmid	human
U00023	1	plasmid	human	U00023	U00023	U00023	1	plasmid	human
U00024	1	plasmid	human	U00024	U00024	U00024	1	plasmid	human
U00025	1	plasmid	human	U00025	U00025	U00025	1	plasmid	human
U00026	1	plasmid	human	U00026	U00026	U00026	1	plasmid	human
U00027	1	plasmid	human	U00027	U00027	U00027	1	plasmid	human
U00028	1	plasmid	human	U00028	U00028	U00028	1	plasmid	human
U00029	1	plasmid	human	U00029	U00029	U00029	1	plasmid	human
U00030	1	plasmid	human	U00030	U00030	U00030	1	plasmid	human
U00031	1	plasmid	human	U00031	U00031	U00031	1	plasmid	human
U00032	1	plasmid	human	U00032	U00032	U00032	1	plasmid	human
U00033	1	plasmid	human	U00033	U00033	U00033	1	plasmid	human
U00034	1	plasmid	human	U00034	U00034	U00034	1	plasmid	human
U00035	1	plasmid	human	U00035	U00035	U00035	1	plasmid	human
U00036	1	plasmid	human	U00036	U00036	U00036	1	plasmid	human
U00037	1	plasmid	human	U00037	U00037	U00037	1	plasmid	human
U00038	1	plasmid	human	U00038	U00038	U00038	1	plasmid	human
U00039	1	plasmid	human	U00039	U00039	U00039	1	plasmid	human
U00040	1	plasmid	human	U00040	U00040</				

REFERENCE	TITLE	ORGANISM
1 (bases 1 to 644)	Homo sapiens	Homo sapiens
Jia, L.B., Young, M.F., Touchman, J.W., Bouffard, G.G., Beckstrom-Sternberg, S.M., Green, E.D., Powell, J.I., Yang, L.M., Robey, P.G., Holcomb, R.N. and Francomano, C.A.	Human Genome Research Institute	Human Genome Research Institute
SGAP: The Skeletal Anatomy Project	Medical Genetics Branch	Medical Genetics Branch
Unpublished (1997)	Contact: Libin Jia	Contact: Libin Jia
	National Human Genome Research Institute	National Human Genome Research Institute
	10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA	10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
	Tel: 301-496-7157	Tel: 301-496-7157
	Fax: 301-496-7157	Fax: 301-496-7157
	Email: libin@nih.gov	Email: libin@nih.gov
	DNA Sequencing and analyses by National Institutes of Health	DNA Sequencing and analyses by National Institutes of Health
	Intramural Sequencing Center, NIDCR	Intramural Sequencing Center, NIDCR
	Plate: 27 row: C column: 11	Plate: 27 row: C column: 11
	Seq primer: 21M13 forward primer (ABI).	Seq primer: 21M13 forward primer (ABI).
	Location/Qualifiers	Location/Qualifiers
	1. 644	1. 644
	/organism="Homo sapiens"	/organism="Homo sapiens"
	/db_xref="taxon:9606"	/db_xref="taxon:9606"
	/clone="NHTC-CN27C11"	/clone="NHTC-CN27C11"
	/clone_id="Normal Human Trabecular Bone Cells"	/clone_id="Normal Human Trabecular Bone Cells"
	/sex="Female"	/sex="Female"
	/tissue_type="Bone"	/tissue_type="Bone"
	/cell_type="Trabecular Bone Cells"	/cell_type="Trabecular Bone Cells"
	/lab_host="SURE"	/lab_host="SURE"
	/note="Organ: Hip; Vector: pBluescript; Site: 1; Ecoli; Library constructed by Dr. Marian Young and Dr. Pamela Gehron Robey (NIDCR)"	/note="Organ: Hip; Vector: pBluescript; Site: 1; Ecoli; Library constructed by Dr. Marian Young and Dr. Pamela Gehron Robey (NIDCR)"
	Gehron Robey (NIDCR)"	Gehron Robey (NIDCR)"
	146 a 195 c 154 g 149 t	146 a 195 c 154 g 149 t
	BASE COUNT	BASE COUNT
	ORIGIN	ORIGIN
	Query Match	Query Match
	Best Local Similarity 100.0%; Pred. NO. 8.8e-163; Indels 0; Gaps 0;	Best Local Similarity 100.0%; Pred. NO. 8.8e-163; Indels 0; Gaps 0;
	Matches 636; Conservative 0; Mismatches 0;	Matches 636; Conservative 0; Mismatches 0;
	134 CCGGCAAAACCTGTGTGATCGAGGGCCCTACTCGAACCCCTACTGACCCCTACTGCG 193	134 CCGGCAAAACCTGTGTGATCGAGGGCCCTACTCGAACCCCTACTGACCCCTACTGCG 193
	9 CCGGCAAAACCTGTGTGATCGAGGGCCCTACTCGAACCCCTACTGACCCCTACTGCG 68	9 CCGGCAAAACCTGTGTGATCGAGGGCCCTACTCGAACCCCTACTGACCCCTACTGCG 68
	194 GTCCGTACCCAGCAGCTGCCCCACCACTCTCAGCTCCAACTATCCAGATCTCCAGCG 253	194 GTCCGTACCCAGCAGCTGCCCCACCACTCTCAGCTCCAACTATCCAGATCTCCAGCG 253
	69 GTCCGTACCCAGCAGCTGCCCCACCACTCTCAGCTCCAACTATCCAGATCTCCAGCG 128	69 GTCCGTACCCAGCAGCTGCCCCACCACTCTCAGCTCCAACTATCCAGATCTCCAGCG 128
	254 CTCTTATATGCGGTTTGATACAGATGATGATGAAGCAACCAATGTGTGATGAGCG 313	254 CTCTTATATGCGGTTTGATACAGATGATGATGAAGCAACCAATGTGTGATGAGCG 313
	129 CTCTTATATGCGGTTTGATACAGATGATGATGAAGCAACCAATGTGTGATGAGCG 188	129 CTCTTATATGCGGTTTGATACAGATGATGATGAAGCAACCAATGTGTGATGAGCG 188
	314 AGTGTCAACAGATTTCCACAGTGCACACCCACCAAGATGATGATGATGATGATGAGCG 248	314 AGTGTCAACAGATTTCCACAGTGCACACCCACCAAGATGATGATGATGATGATGAGCG 248
	189 AGTGTCAACAGATTTCCACAGTGCACACCCACCAAGATGATGATGATGATGATGAGCG 248	189 AGTGTCAACAGATTTCCACAGTGCACACCCACCAAGATGATGATGATGATGATGAGCG 248
	374 GGTAAACCTGCTCCAGCAGCAGATTTGGCTTTGGAAGGCCAGTGTAGACATG 433	374 GGTAAACCTGCTCCAGCAGCAGATTTGGCTTTGGAAGGCCAGTGTAGACATG 433
	249 GGTAAACCTGCTCCAGCAGCAGATTTGGCTTTGGAAGGCCAGTGTAGACATG 308	249 GGTAAACCTGCTCCAGCAGCAGATTTGGCTTTGGAAGGCCAGTGTAGACATG 308
	434 ATGATGTGCTATGATGTTACTGCGACGACCTCTGTGCGAATGTTCTGTGATCTTCTT 493	434 ATGATGTGCTATGATGTTACTGCGACGACCTCTGTGCGAATGTTCTGTGATCTTCTT 493
	309 ATGATGTGCTATGATGTTACTGCGACGACCTCTGTGCGAATGTTCTGTGATCTTCTT 368	309 ATGATGTGCTATGATGTTACTGCGACGACCTCTGTGCGAATGTTCTGTGATCTTCTT 368
	494 GTACATGCAACCTGTTTACCTCATAGATGATGATGATGATGATGATGATGATGATG 428	494 GTACATGCAACCTGTTTACCTCATAGATGATGATGATGATGATGATGATGATGATG 428
	369 GTACATGCAACCTGTTTACCTCATAGATGATGATGATGATGATGATGATGATGATG 428	369 GTACATGCAACCTGTTTACCTCATAGATGATGATGATGATGATGATGATGATGATG 428
	554 AGTGTGCAACGAGAACCCCTGCTGGAACCTGCTGCAACACCTACGCTCTTCTATCT 613	554 AGTGTGCAACGAGAACCCCTGCTGGAACCTGCTGCAACACCTACGCTCTTCTATCT 613
	429 AGTGTGCAACGAGAACCCCTGCTGGAACCTGCTGCAACACCTACGCTCTTCTATCT 488	429 AGTGTGCAACGAGAACCCCTGCTGGAACCTGCTGCAACACCTACGCTCTTCTATCT 488
	614 GCCCGTGTGACCCAGATATGACTTGAGGAAGATGGCGTTCAATTGAGATATGAGCG 673	614 GCCCGTGTGACCCAGATATGACTTGAGGAAGATGGCGTTCAATTGAGATATGAGCG 673

[illegible]

```

OY 431 TTGATGAATGCTGCTATGTTACTGTCACAGCTGCTGCGAATGTTCTGATCTTAT 490
    |||||||
Db 181 TTGATGAATGCTGCTATGTTACTGTCACAGCTGCTGCGAATGTTCTGATCTTAT 490
OY 491 CTGTATGATGCAACCTGTTTACCTCAATGAGATGGAAGTCTTGGCAAGATGGA 240
    |||||||
Db 241 CCTGTATGCAACCTGTTTACCTCAATGAGATGGAAGTCTTGGCAAGATGGA 240
OY 551 ACGATGTCACAGCAACCTGCTGTCGCAACCTGCTGTCACACCTGAGGCTTTTCA 300
    |||||||
Db 301 ACGATGTCACAGCAACCTGCTGTCGCAACCTGCTGTCACACCTGAGGCTTTTCA 300
OY 611 TCTGCGCTGTCACAGCAACCTGCTGTCGCAACCTGCTGTCACACCTGAGGCTTTTCA 360
    |||||||
Db 361 TCTGCGCTGTCACAGCAACCTGCTGTCGCAACCTGCTGTCACACCTGAGGCTTTTCA 360
OY 671 ACGATGTCACAGCTTCTGAGTCTGCTGTCGCAACCTGTCGCAACCTGAGGCTTTTCA 420
    |||||||
Db 421 ACGATGTCACAGCTTCTGAGTCTGCTGTCGCAACCTGTCGCAACCTGAGGCTTTTCA 420
OY 731 ACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
    |||||||
Db 481 ACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
OY 791 TCACAGATGTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 540
    |||||||
Db 541 TCACAGATGTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 540
OY 851 AAGGGGGCTTCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
    |||||||
Db 601 AAGGGGGCTTCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
OY 911 ATAAACCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
    |||||||
Db 661 ATAAACCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660

RESULT 10
LOCUS B0716088 933 bp mRNA linear EST 16-JUL-2002
DEFINITION AGENCOURT_8099803 Lupski_sympathetic_trunk Homo sapiens CDNA clone
IMAGE:6190063 5', mRNA sequence.
B0716088
B0716088.1 GI:21854985
SOURCE EST.
ORGANISM human.
REFERENCE Homo sapiens
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
JOURNAL NIH-MGC http://mgi.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs.fda.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13588 row: e column: 08
High quality sequence stop: 501.
Location/Qualifiers
1..933
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="6190063"
/clone_lib="Lupski_sympathetic_trunk"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"

```

```

/lab_host="DH10B"
/Note="Vector: PCW-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCCACCCAGCGGCG-3' and
5'-GACTAGCTTGTAGATCGGAGCGGCGGCGCTT-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."
BASE COUNT 225 a 284 c 245 g 179 t
ORIGIN

```

```

Query Match 45.4%; Score 576.4; DB 14; Length 933;
Best Local Similarity 94.6%; Pred. No. 2e-146;
Matches 619; Conservative 0; Mismatches 31; Indels 4; Gaps 2;

```

```

OY 527 ATGAAGGCTTGGCAAGATGTAAGAGAGTGTGCAACCAACCTGCTGCAACCT 586
    |||||||
Db 2 ATGAAGGCTTGGCAAGATGTAAGAGAGTGTGCAACCAACCTGCTGCAACCT 586
OY 587 GCGTCAACCTGAGGCTCTTTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 61
    |||||||
Db 62 GCGTCAACCTGAGGCTCTTTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 61
OY 647 ATGCGCTTCAATTCAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 121
    |||||||
Db 122 ATGCGCTTCAATTCAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 121
OY 707 AGTGTGTGAACAGCGCGGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 181
    |||||||
Db 182 AGTGTGTGAACAGCGCGGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 181
OY 767 ATGACACCGAAGCTGCGCAAGATGATGATGATGATGATGATGATGATGATGATGAT 241
    |||||||
Db 242 ATGACACCGAAGCTGCGCAAGATGATGATGATGATGATGATGATGATGATGATGAT 241
OY 827 TCACAGCAGCTGCTTACATTTTACAGAGGGGCTTCAATGATGATGATGATGATGAT 301
    |||||||
Db 302 TCACAGCAGCTGCTTACATTTTACAGAGGGGCTTCAATGATGATGATGATGATGAT 301
OY 887 AGGAGCTTATCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 361
    |||||||
Db 362 AGGAGCTTATCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 361
OY 947 GCAGAGACAGCGCTTACCATCTTGTACCGGACATGAGCTGCTGCTGCTGCTGCT 421
    |||||||
Db 422 GCAGAGACAGCGCTTACCATCTTGTACCGGACATGAGCTGCTGCTGCTGCTGCT 421
OY 1007 TTCCCGCTGACATCTTCCAAATGCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 481
    |||||||
Db 482 TTCCCGCTGACATCTTCCAAATGCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 481
OY 1067 TCCAGATCAATCTGGGAATGA-GGGCAGAGAATTTTTCATGCGGCAAA--CGGGCCC 541
    |||||||
Db 542 TCCAGATCAATCTGGGAATGAAGGGGAGAGAGATTTTTCATGCGGCAAA--CGGGCCC 541
OY 1123 ATCAGTGCACCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1176
    |||||||
Db 602 ATCAGTGCACCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1176

```

```

RESULT 11
LOCUS BF090699/c 627 bp mRNA linear EST 19-OCT-2000
DEFINITION OY2_NT0048-120900-360-all NT0048 Homo sapiens CDNA, mRNA sequence.
ACCESSION BF090699
VERSION BF090699.1 GI:10896409
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE	AUTHORS	TITLE	JOURNAL MEDLINE COMMENT
1 (bases 1 to 627)	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.		
	Dias Neto, E., Garcia Correa, R., Veljovski-Almeida, S., Brites, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
	2020263		
	Contact: Simpson A.J.G.		
	Laboratory of Cancer Genetics		
	Ludwig Institute for Cancer Research		
	Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil		
	Tel.: +55-11-2704922		
	Fax: +55-11-2707001		
	Email: asimpson@ludwig.org.br		
	This sequence was derived from the FAPESP/LICR Human Cancer Genome Project.		
	This sequence was derived from the following URL		
	(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=8t2-QV2-NT0048-120)		
	900-360-allec3-2000-09-12c4-1)		
	Seq primer: puc 18 forward		
	High quality sequence start: 10		
	High quality sequence stop: 606.		
FEATURES			
source			
	1..627		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone_lib="NT0048"		
	/dev_stage="Adult"		
	/note="Organ: nervous tumor; Vector: puc18; Site 1: Smal; Site 2: Smal; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."		
BASE COUNT	148 a 161 c 170 g 148 t		
ORIGIN			
Query Match	45.2%; Score 573.6; DB 12; Length 627;		
Best Local Similarity	98.2%; Pred. No. 9.7e-146;		
Matches 591; Conservative	0; Mismatches 9; Indels 2; Gaps 1;		
369	AGCGGGATACACCTGCTCTCCACGAGATATGCGTTGGAAGCCAGTCTTGA 428		
613	AAGCGGGACACCTGCTCTCCACGAGATATGCGTTGGAAGCCAGTCTTGA 554		
429	CATTGATGATGTCGCTATGTTACTGCCAGCAGCTGTGCGAATGTTCCGATCTA 488		
553	CATTGATGATGTCGCTATGTTACTGCCAGCAGCTGTGCGAATGTTCCGATCTA 494		
489	TTCTTGATGATGATGTCGCTATGTTACTGCCAGCAGCTGTGCGAATGTTCCGATCTA 548		
493	TTCTTGATGATGATGTCGCTATGTTACTGCCAGCAGCTGTGCGAATGTTCCGATCTA 548		
549	GAAGAGTGTGCCACCGAAGACCCCTGCGTCAAACTGCGTCAACACCTAGGCTCTTT 608		
433	GAAGAGTGTGCCACCGAAGACCCCTGCGTCAAACTGCGTCAACACCTAGGCTCTTT 608		
609	CATCGCCCGCTGACCCAGATATGAACTTGAGAAATAGCGCTTCAATGACAGATAT 668		
373	CATCGCCCGCTGACCCAGATATGAACTTGAGAAATAGCGCTTCAATGACAGATAT 668		
669	GGAGAGTGTGCCACCGAAGACCCCTGCGTCAAACTGCGTCAACACCTAGGCTCTTT 728		
313	GGAGAGTGTGCCACCGAAGACCCCTGCGTCAAACTGCGTCAACACCTAGGCTCTTT 728		
729	ATTCTTGTCTCTGCTCCCTCCAGGCTACATCTCTGATGATCAACGAGTGTCCAA 788		

Db	253	ATACTTGTGCTCCCTGCCTCCAGGCTACATCTGCTGGATACAAACCACTGCAAGA	194
Qy	789	CATCAACGAATGTATAGACACAGAACCCACAGCTGCAACCTGGACAGAGCTGCTACAAATT	848
Db	193	CATCAACGAATGTATAGACACAGAACCCACAGCTGCAACCTGGACAGAGCTGCTACAAATT	134
Qy	849	ACAAGGGGGGCTTCCAAATGCATCGAACCCCATCCGCTGTGAGGAGCCCTTATCTGAGAGTCAG	908
Db	133	ACAAGGGGGGCTTCCAAATGCATCGAACCCCATCCGCTGTGAGGAGCCCTTATCTGAGAGTCAG	74
Qy	909	TGATTAACCGCTGTATGTCTCTGCTGAGAACCCCTGCTGTCGACAGACCAAGCCCTTACCAT	968
Db	73	TGATTAACCGCTGTATGTCTCTGCTGAGAACCCCTGCTGTCGACAGACCAAGCCCTTACCAT	16
Qy	969	CT 970	
Db	15	CT 14	
RESULT 12			
LOCUS	B1825121	882 bp	mRNA linear EST 04-OCT-2001
DEFINITION	603038662p1 NIH_MGC_115 Homo sapiens cDNA, clone IMAGE:5179550 5',		
ACCESSION	B1825121		
VERSION	B1825121.1	GI:15936671	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	NIH-MGC http://mgc.ncl.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cga@bbs-remail.nih.gov		
	Tissue Procurement: Life Technologies, Inc.		
	cDNA Library Preparation: Life Technologies, Inc.		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/MLN at:		
	http://image.llnl.gov		
	Plate: LHAM11447 row: 1 column: 15		
	High quality sequence start: 14		
	High quality sequence stop: 844.		
FEATURES	Location/Qualifiers		
source	1..882		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone_image="5179550"		
	/clone_id="NIH_MGC_115"		
	/lab_host="DH10B"		
	/note="Organ: pooled brain, lung, testis; Vector:		
	pcmv-SpO6; Site:1: NotI; Site:2: EcoRV (destroyed); RNA		
	source anonymous pool of 6 male brains, age range 23-27; 1		
	male lung, age 27; and 1 male testis, age 69. Library is		
	oligo-dT primed and directionally cloned (EcoRV site is		
	destroyed upon cloning). Average insert size 1.8 kb,		
	insert size range 1-3 kb. Library is normalized and		
	enriched for full-length clones and was constructed by C		
	Gruber (Invitrogen). Research Genetics tracking code		
	021. Note: this is a NIH_MGC Library."		
BASE COUNT	191 a 277 c 216 g 197 t	1 others	
ORIGIN			
Query Match	44.6%;	Score 566.2;	DB 13; Length 882;
Best Local Similarity	97.9%;	Pred. NO. 1.2e-143;	
Matches	615; Conservative	0; Mismatches	9; Indels 4; Gaps 4;
1	CAGTGCACGAATGGCTTGACCTGGATGCGCCAGTACAGACAGTGTTTAGATTTGATGAA	60	

```

Db      254 CAGTGCAGCAATGCGCTTTACCTGGATGCCCACTCAGACAGTGTAGATTGATGAA 313
QY      61 TCGCGAACCATCCCGAGGCGCTGCGGAGAGAGACATGATGTGTAAACAAATGCGGG 120
Db      314 TGCCGACCAATCCCGAGGCGCTGCGGAGAGACATGATGTGTAAACAAATGCGGG 373
QY      121 TATTATGATTCCTCCCGGACAAACCTGTGTATCGAGGGCCCTACTGGAACCCCTACTCG 180
Db      374 TATTATGATTCCTCCCGGACAAACCTGTGTATCGAGGGCCCTACTGGAACCCCTACTCG 433
QY      181 ACCCCTACTCAGTCCGTACCCAGCAGCGCCCACTCTCAGCTCCAACTATGCC 240
Db      434 ACCCCTACTCAGTCCGTACCCAGCAGCGCCCACTCTCAGCTCCAACTATGCC 493
QY      241 ACGATCTCCAGGCGCTTATATGCGGCTTTGATACCATGATGATGAAGCAACCAATGT 300
Db      494 ACGATCTCCAGGCGCTTATATGCGGCTTTGATACCATGATGATGAAGCAACCAATGT 553
QY      301 GGTGATGTGAGAGAGTGCACAGATTCGCCACAGTGCACACCCCACTGATGCGATC 360
Db      554 GTGATGTGAGAGAGTGCACAGATTCGCCACAGTGCACACCCCACTGATGCGATC 613
QY      361 AATACTGAAGGCGGTACACCTGCTCTGACACGCGAGATTTGGCTTGGAAAGCCAG 420
Db      614 AATACTGAAGGCGGTACACCTGCTCTGACACGCGAGATTTGGCTTGGAAAGCCAG 673
QY      421 TGCTTAGACATGATGATGCTGCTGATGCTTACTGCGAGAGCTGTGGCAATGCTCT 480
Db      674 TGCTTAGACATGATGATGCTGCTGATGCTTACTGCGAGAGCTGTGGCAATGCTCT 733
QY      481 GGATCTTATCTTGTACATGCAACCCCT-GGTTTACCTCAATGAGATGG-AGGTCCT 538
Db      734 GGATCTTATCTTGTACATGCAACCCCTGGGTTTACCTCAATGAGATGGAGGTCCT 793
QY      539 GCCAAGATGTGAAG-AGTGTGCCACCGAGAACCCCTGCGTGCACAACTGCGTCAACAC 597
Db      794 GCCAAGATGTGAAGAGTGTGTCACCGAGAACCCCTGCGTGCACAACTGCGTCAACAC 853
QY      598 TACGGCTCTTTCATCGCGCGCTGTGACC 625
Db      854 TA-GGGCTTTTCATCGCGCGCTGTGACC 880

RESULT 13
B0930927 870 bp mRNA linear EST 20-AUG-2002
LOCUS AGENCOUTF_8955548 NCL.CGAP_Mam2 Mus musculus cDNA clone
DEFINITION IMAGE:6439732 5', mRNA sequence.
ACCESSION B0930927.1 GI:22345958
VERSION EST.
KEYWORDS Mus musculus
SOURCE house mouse
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 870)
NIH-MGC http://mgc.ncl.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13960 row: d column: 05
High quality sequence stop: 651.
Location/Qualifiers
1..870
/organism="Mus musculus"

```

```

Query Match 44.5% Score 565.2; DB 14; Length 870;
Best Local Similarity 87.9% Pred. No. 2.3e-143;
Matches 649; Conservative 0; Mismatches 85; Indels 4; Gaps 3;

Db      1 CAGTGCAGCAATGCGCTTTGACCTGATCGCCAGTCCAGACAGTGTAGATTGATGAA 60
QY      123 CAGTGCAGCAATGCGCTTTGACCTGATCGCCAGTCCAGACAGTGTAGATTGATGAA 182
Db      61 TCGCGAACCATCCCGAGGCGCTGCGGAGAGACATGATGTGTAAACAAATGCGGG 120
QY      183 TGCCGACCAATCCCGAGGCGCTGCGGAGAGACATGATGTGTAAACAAATGCGGG 242
Db      121 TATTATGATTCCTCCCGGACAAACCTGTGTATCGAGGGCCCTACTGGAACCCCTACTCG 180
QY      243 TATTATGATTCCTCCCGGACAAACCTGTGTATCGAGGGCCCTACTGGAACCCCTACTCG 302
Db      181 ACCCCTACTCAGTCCGTACCCAGCAGCGCCCACTCTCAGCTCCAACTATGCC 240
QY      303 ACATCTACTCAGGCGCTTATATGCGGCTTTGATACCATGATGATGAAGCAACCAATGT 362
Db      241 ACGATCTCCAGGCGCTTATATGCGGCTTTGATACCATGATGATGAAGCAACCAATGT 538
QY      363 ACATCTACTCAGGCGCTTATATGCGGCTTTGATACCATGATGATGAAGCAACCAATGT 793
Db      301 GGTGATGTGAGAGAGTGCACAGATTCGCCACAGTGCACACCCCACTGATGCGATC 360
QY      423 GTGATGTGAGAGAGTGCACAGATTCGCCACAGTGCACACCCCACTGATGCGATC 482
Db      361 AATACTGAAGGCGGTACACCTGCTCTGACACGCGAGATTTGGCTTGGAAAGCCAG 420
QY      483 AACACTGAGAGAGTGTACACTGCTGCGACCGATGGTACTGCTTGGAAAGGCGAG 542
Db      421 TGCTTAGACATGATGATGCTGCTGATGCTTACTGCGAGAGCTGTGGCAATGCTCT 480
QY      543 TGCTTAGACATGATGATGCTGCTGATGCTTACTGCGAGAGCTGTGGCAATGCTCT 602
Db      481 GGATCTTATCTTGTACATGCAACCCCTGGTTTTACCTCAATGAGATGGAGGTCCTGC 540
QY      603 GGATCTTATCTTGTACATGCAACCCCTGGTTTTACCTCAATGAGATGGAGGTCCTGC 662
Db      541 CAAGATGTGAAGAGTGTGCGACCGAGAACCCCTGCGTGCACAACTGCGTCAACAC 600
QY      663 CAAGATGTGAAGAGTGTGCGACCGAGAACCCCTGCGTGCACAACTGCGTCAACAC 722
Db      601 GGCTCTTCAATCTGCGCTGTGACCCAGAGATATG-ACTTGAAGAGATGGCGTCA-TT 658
QY      723 GGCTCTTCAATCTGCGCTGTGACCCAGAGATATGAACTTGAAGAGATGGCGTCA-TT 782
Db      659 GCAATGATGAGAGAGTGCAGC--TTCGTGAGATTCCTCTGCGCAACATGATGTGTAA 716
QY      783 GCAATGATGAGAGAGTGCAGCCTTCGAGAGTTCCTCTGCGCAACATGATGTGTAA 842
Db      717 CCAAGCCCGGACATACCT 734
QY      843 ACAAGCCGGGCTCATACCT 860

RESULT 14
B1768516

```

LOCUS BI768516 912 bp mRNA linear EST 25-SEP-2001
 DEFINITION 603057118F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5206588 5',
 mRNA sequence.
 ACCESSION BI768516
 VERSION BI768516.1 GI:15760094
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 912)
 AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLM1518 row: C column: 05
 High quality sequence stop: 810.
 Location/Qualifiers
 1. 912
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5206588"
 /clone_1lb="NIH_MGC_122"
 /lab_host="DH10B"
 /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
 Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
 anonymous pool of 24 week female lung, 16 week female
 spleen, and 20-22 week male spleens. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.4 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 026. Note:
 this is a NIH_MGC Library."
 BASE COUNT 230 a 262 c 225 g 195 t
 ORIGIN
 Query Match 44.5%; Score 564.4; DB 13; Length 912;
 Best Local Similarity 99.8%; Pred. No. 3.8e-143;
 Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 704 ATGAGTGTGACACAGCCGCGACATACCTCTGCTCGCCCTCCAGGCTACATCCTGC 763
 |||
 Db 1 ATGAGTGTGACACAGCCGCGACATACCTCTGCTCGCCCTCCAGGCTACATCCTGC 60
 Oy 764 TGTATGACAAACGAGTGTGCAAGATCAAGATGTAGAGCAGAGAACACAGCTGCA 823
 |||
 Db 61 TGTATGACAAACGAGTGTGCAAGATCAAGATGTAGAGCAGAGAACACAGCTGCA 120
 Oy 824 ACCTGCACAGACGCTGCTACATTTTCAAGGGGGCTTCAATGCAATGCAATCCATCCGCT 883
 |||
 Db 121 ACCTGCACAGACGCTGCTACATTTTCAAGGGGGCTTCAATGCAATGCAATCCATCCGCT 180
 Oy 884 GTGAGAGCCCTTATGTAGAGATCAGTAAACCGGTGATGCTGCTGAGAAACCTTG 943
 |||
 Db 181 GTGAGAGCCCTTATGTAGAGATCAGTAAACCGGTGATGCTGCTGAGAAACCTTG 240
 Oy 944 GCTCAGAGACAGCCCTTTTACCATCTGTACCGGGAGATGAGAGTGTGAGAGAGCT 1003
 |||
 Db 241 GCTCAGAGACAGCCCTTTTACCATCTGTACCGGGAGATGAGAGTGTGAGAGAGCT 300
 Oy 1004 CCGTTCCCGCTGACATCTTCAATGCAAGCAAGACCGGCTACCTGGGGCTATTACA 1063
 |||
 Db 301 CCGTTCCCGCTGACATCTTCAATGCAAGCAAGACCGGCTACCTGGGGCTATTACA 360

Oy 1064 TTTTCAGATCAATCTGGGAATGAGGCGAGAAATTTTACATGCGGCAACGGGCCCA 1123
 |||
 Db 361 TTTTCAGATCAATCTGGGAATGAGGCGAGAAATTTTACATGCGGCAACGGGCCCA 420
 Oy 1124 TCAGTGCCACCCCTGTGTAGACACGGCCCATCAAGAGGCCCGGGAATCCAGCTGACT 1183
 |||
 Db 421 TCAGTGCCACCCCTGTGTAGACACGGCCCATCAAGAGGCCCGGGAATCCAGCTGACT 480
 Oy 1184 TGGAAATGATCTACTGTCACACACTGTATCAATCTTACAGAGCAGCTCGTATCCGACTGC 1243
 |||
 Db 481 TGGAAATGATCTACTGTCACACACTGTATCAATCTTACAGAGCAGCTCGTATCCGACTGC 540
 Oy 1244 GGATATATGTATCGCAGTACCCATTTC 1269
 |||
 Db 541 GGATATATGTATCGCAGTACCCATTTC 566
 RESULT 15
 LOCUS BI764020 880 bp mRNA linear EST 25-SEP-2001
 DEFINITION 603043221F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5183693 5',
 mRNA sequence.
 ACCESSION BI764020
 VERSION BI764020.1 GI:15755598
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 880)
 AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLM1458 row: 1 column: 06
 High quality sequence stop: 846.
 Location/Qualifiers
 1. 880
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5183693"
 /clone_1lb="NIH_MGC_116"
 /lab_host="DH10B"
 /note="Organ: pooled colon, kidney, stomach; Vector:
 pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of 3 colons, age 26 yo male, 49 yo
 female, 71 yo male colon; and 70 yo female kidney, and pool of 2
 stomachs, 62 yo male and 40 yo female. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.4 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 023. Note: this is a NIH_MGC Library."
 BASE COUNT 192 a 278 c 215 g 195 t
 ORIGIN
 Query Match 43.3%; Score 549.6; DB 13; Length 880;
 Best Local Similarity 98.3%; Pred. No. 4.1e-139;
 Matches 619; Conservative 0; Mismatches 4; Indels 7; Gaps 6;
 Oy 1 CAGTGCAGAAATGGCTTTGACCTGATGCCAGTACAGACAGTCTTTAGATATTGATGAA 60
 |||
 Db 252 CAGTGCAGAAATGGCTTTGACCTGATGCCAGTACAGACAGTCTTTAGATATTGATGAA 311

```
OY 61 TCGGACCATCCCGAGGCTGCCGAGGAGACATGATGTGTATACCAAAATGCGGG 120
    |||||||
Db 312 TCGGACCATCCCGAGGCTGCCGAGGAGACATGATGTGTATACCAAAATGCGGG 371
    |||||||
OY 121 TATTTATGCAATTCGCCGAGCAAAACCTGTGTATGAGAGGCTTACTGAACCCCTACTCG 180
    |||||||
Db 372 TATTTATGCAATTCGCCGAGCAAAACCTGTGTATGAGAGGCTTACTGAACCCCTACTCG 431
    |||||||
OY 181 ACCCCCTACTGAGTCCGTACCCAGAGCTGCCACACTCTCAGCTCCAAACTATGCC 240
    |||||||
Db 432 ACCCCCTACTGAGTCCGTACCCAGAGCTGCCACACTCTCAGCTCCAAACTATGCC 491
    |||||||
OY 241 ACGATCTCCAGGCTCTTATATGCGCTTGTGATACCAATGATGATGAAGCAACCAATGT 300
    |||||||
Db 492 ACGATCTCCAGGCTCTTATATGCGCTTGTGATACCAATGATGATGAAGCAACCAATGT 551
    |||||||
OY 301 GTGATGTGACAGAGTGTGCAACAGATTGCCACAGTGCACACCCAGATCTGCATC 360
    |||||||
Db 552 GTGATGTGACAGAGTGTGCAACAGATTGCCACAGTGCACACCCAGATCTGCATC 611
    |||||||
OY 361 AATACTGAAGCGGCTACACCTGCTCCTGCAACGAGGATATGGCTTCTGGAAGGCCAG 420
    |||||||
Db 612 AATACTGAAGCGGCTACACCTGCTCCTGCAACGAGGATATGGCTTCTGGAAGGCCAG 671
    |||||||
OY 421 TGCTTAGACATTTGATGATGTGCTATGCTTACTGCTGCAAGCA -GCTCTGAGGATGTTCC 479
    |||||||
Db 672 TGCTTAGACATTTGATGATGTGCTATGCTTACTGCTGCAAGCAAGCTCTGTCGAATGTTCC 731
    |||||||
OY 480 TGATCTTATTTCTTTGA -CATGCAACCCCTGTTTACCTCAATGAGGATGAA -GGTC 536
    |||||||
Db 732 TGATCTTATTTCTTTGA -CATGCAACCCCTGTTTACCTCAATGAGGATGAA -GGTC 791
    |||||||
OY 537 TTGCCAAGA -TGTCACGAGTGTGCCACCGAGAACCCCTGCGTGCACAACTGCGTCAACA 595
    |||||||
Db 792 TTGCCAAGATTTGACGAGTGTGCCACCGAGAA -CCCTGCGTCAAA -CTGCGTCAACA 849
    |||||||
OY 596 CCTACGGCTCTTTCATCTGCGCGCTGTGACC 625
    |||||||
Db 850 CCTACGGATCTTTCACCTGCGCGCTGTGACC 879
    |||||||
```

Search completed: July 3, 2003, 17:36:08
Job time : 1873.78 secs

QY	483	CACGATCTCCAGGCGCTCTTATATGCGCTTTGGATACCAAGATGATGAACCAACCAATG	542
Db	764	CACGATCTCCAGGCGCTCTTATATGCGCTTTGGATACCAAGATGATGAACCAACCAATG	823
QY	543	TGTGGATGTGACGAGTGTGCACAGATATCCACAGTGCACAAACCCCAACCGATCTGCAT	602
Db	824	TGTGGATGTGACGAGTGTGCACAGATATCCACAGTGCACAAACCCCAACCGATCTGCAT	883
QY	603	CAATATCTGAAGCGGGGTACACCTGCTCCTGCACGACGAGATATGGCTTCGGAAGGCCA	662
Db	884	CAATATCTGAAGCGGGGTACACCTGCTCCTGCACGACGAGATATGGCTTCGGAAGGCCA	943
QY	663	GTGCTTGAACATTTATGAATGTCCTATAGTTACTGTGCACAGAGCTGTGTGGATATGTTCC	722
Db	944	GTGCTTGAACATTTATGAATGTCCTATAGTTACTGTGCACAGAGCTGTGTGGATATGTTCC	1003
QY	723	TGATCTCTATCTTGTATGATGCACACCGCTGTATTTACCTCAATAGATAGGAAGTCTGTG	782
Db	1004	TGATCTCTATCTTGTATGATGCACACCGCTGTATTTACCTCAATAGATAGGAAGTCTGTG	1063
QY	783	CCAAGATGTGAACGAGTGTGCCACCCGAGAACCCCTGGGTGCAAACCTGTGCTCAACCTTA	842
Db	1064	CCAAGATGTGAACGAGTGTGCCACCCGAGAACCCCTGGGTGCAAACCTGTGCTCAACCTTA	1123
QY	843	CGGCTCTTTCATCTGCGGCTGTGACCCGAGATATGAACCTTAGGAAGATGGCGTTCAATG	902
Db	1124	CGGCTCTTTCATCTGCGGCTGTGACCCGAGATATGAACCTTAGGAAGATGGCGTTCAATG	1183
QY	903	CAGTGATATGAGCAGATGACAGCTTCTCTGAATTCCTCTGCCAACATGATGTGTGAACCA	962
Db	1184	CAGTGATATGAGCAGATGACAGCTTCTCTGAATTCCTCTGCCAACATGATGTGTGAACCA	1243
QY	963	GCCGGGACATATCTTGCCTCTGCTCCCTCCAGGCTTACATCCTGCTGGATGTGAACCCGAG	1022
Db	1244	GCCGGGACATATCTTGCCTCTGCTCCCTCCAGGCTTACATCCTGCTGGATGTGAACCCGAG	1303
QY	1023	CTGGCAAGACATCAACGATGTGACACAGAGAACACACACGTGCACACCTGCACGACAGCTG	1082
Db	1304	CTGGCAAGACATCAACGATGTGACACAGAGAACACACACGTGCACACCTGCACGACAGCTG	1363
QY	1083	CTACAAATTTACAAAGGGGCTTCAAAATGATGACACCCATCCGCTGTAGAGAGCCTTATCT	1142
Db	1364	CTACAAATTTACAAAGGGGCTTCAAAATGATGACACCCATCCGCTGTAGAGAGCCTTATCT	1423
QY	1143	GAGATCATGATATACCGCTGTATGTGCTCCTGTGAGAAACCTGTGCTCAAGACACAGCC	1202
Db	1424	GAGATCATGATATACCGCTGTATGTGCTCCTGTGAGAAACCTGTGCTCAAGACACAGCC	1483
QY	1203	CTTTACCATCTTGTACCGGGGACATGAGAGTGTGTCAAGACGCTCCGTTCCCGCTGACAT	1262
Db	1484	CTTTACCATCTTGTACCGGGGACATGAGAGTGTGTCAAGACGCTCCGTTCCCGCTGACAT	1543
QY	1263	CTTCCAAATGCAACCCAGACCCGCTACCCCTGGGGCCATATACATTTTCCAGATCAATC	1322
Db	1544	CTTCCAAATGCAACCCAGACCCGCTACCCCTGGGGCCATATACATTTTCCAGATCAATC	1603
QY	1323	TGGGATATGAGGGCAGAGAAATTTTACATGCGGCAAAAGGGGCCCATCACTGCTCACCTGGT	1382
Db	1604	TGGGATATGAGGGCAGAGAAATTTTACATGCGGCAAAAGGGGCCCATCACTGCTCACCTGGT	1663
QY	1383	GATGACACGCCCCCATCAAAAGGGGCCCGGGAATCCAGCTGGAATTTGGAAATGATCACCTGT	1442
Db	1664	GATGACACGCCCCCATCAAAAGGGGCCCGGGAATCCAGCTGGAATTTGGAAATGATCACCTGT	1723
QY	1443	CAACACTGTCTCAACTTTCAGAGGACAGCTCCGATATCCATGCACTGGGGAATATATGTGTGCA	1502
Db	1724	CAACACTGTCTCAACTTTCAGAGGACAGCTCCGATATCCATGCACTGGGGAATATATGTGTGCA	1783
QY	1503	GTACCCATTTGTGAGAGCTGTGGGCTGTGAGGCTTCCAGACGCTGCTCTCATTTGGCACCAAGGA	1562
Db	1784	GTACCCATTTGTGAGAGCTGTGGGCTGTGAGGCTTCCAGACGCTGCTCTCATTTGGCACCAAGGA	1843

OY	1563	AGGAGAAAGAGAGAAATTAACAGAGAGAAATGAGAGCAGACAGACAGCTTAGCAATTTCTG	1622
Db	1844	CAGAGAGAGAGAGAAATTAACAGAGAGAAATGAGAGCAGACAGACAGCTTAGCAATTTCTG	1903
OY	1623	CTGAAGCTTTCCCGAAGAGTCAACCCCGACTTCTCACTCTCACTCACTGACTATTGCGA	1682
Db	1904	CTGAAGCTTTCCCGAAGAGTCAACCCCGACTTCTCACTCTCACTCACTGACTATTGCGA	1963
OY	1683	CTGTGTCAACCTGCAGAGACTTGCCACCCCGACTTCTATGATACGTTATCAAAAAGTATT	1742
Db	1964	CTGTGTCAACCTGCAGAGACTTGCCACCCCGACTTCTATGATACGTTATCAAAAAGTATT	2023
OY	1743	ATCATTTGTCOCCTCGATAGAAAGATTGTGGGATTTTCAAGGCTTCAGTTATTTTCCA	1802
Db	2024	ATCATTTGTCOCCTCGATAGAAAGATTGTGGGATTTTCAAGGCTTCAGTTATTTTCCA	2083
OY	1803	CTATTTTCAAGAAAAATAGATTAGTTTGGGGGGTGTGAGTCTATGTTCMAAGACTGTG	1862
Db	2084	CTATTTTCAAGAAAAATAGATTAGTTTGGGGGGTGTGAGTCTATGTTCMAAGACTGTG	2143
OY	1863	AACAGCTTGTCTGTACTTCTTCAACCTCTCCACTCTCTCTCTCACTGTGTTACTGCTTGTG	1922
Db	2144	AACAGCTTGTCTGTACTTCTTCAACCTCTCCACTCTCTCTCTCACTGTGTTACTGCTTGTG	2203
OY	1923	CAAGACCCCGGAGACTGGCGGGGAAACCTGGAGATAGTATGTTGGTTTTGGCTAACAA	1982
Db	2204	CAAGACCCCGGAGACTGGCGGGGAAACCTGGAGATAGTATGTTGGTTTTGGCTAACAA	2263
OY	1983	GAGAAAGCTATGTAAACAACACAGAGAGATCCAAAGGTTTTTATAGAAATGTGTTCAA	2042
Db	2264	GAGAAAGCTATGTAAACAACACAGAGAGATCCAAAGGTTTTTATAGAAATGTGTTCAA	2323
OY	2043	AACCATGCTGTGTATTTTCAACCATPAAAAGAGTTTCAGTTGTCTTAAATTTGTATPAC	2102
Db	2324	AACCATGCTGTGTATTTTCAACCATPAAAAGAGTTTCAGTTGTCTTAAATTTGTATPAC	2383
OY	2103	GGTTTAATTCGTCTGTGTCAATTTTGAGTATTTTAAAAAATATGTGGTGAATTTCCCTC	2162
Db	2384	GGTTTAATTCGTCTGTGTCAATTTTGAGTATTTTAAAAAATATGTGGTGAATTTCCCTC	2443
OY	2163	GAAAGGCTTCAGACACATGCTATGTCTGTCTTCCCAAAACCCAGTCTCCTCTCCATATTT	2222
Db	2444	GAAAGGCTTCAGACACATGCTATGTCTGTCTTCCCAAAACCCAGTCTCCTCTCCATATTT	2503
OY	2223	AGCCCAAGTCTTTCTTTGAGAGACCCCTTAATCTTGCTTTCTTTTGAATTTTACCCAATT	2282
Db	2504	AGCCCAAGTCTTTCTTTGAGAGACCCCTTAATCTTGCTTTCTTTTGAATTTTACCCAATT	2563
OY	2283	GGATTGGAAATCAGAGGCTCCAAACGATTAATTAATTTGAAGAGA 2328	
Db	2564	GGATTGGAAATCAGAGGCTCCAAACGATTAATTAATTTGAAGAGA 2609	

Search completed: July 3, 2003, 17:53:38
Job time : 382.284 secs

QY	1143	GAGATCACTGATTAACCCGCTGATGTGTCTCTGCTGAGAAACCCCTGGCTGCAGACACAGCC	1202
Db	1424	GAGGATCACTGATTAACCCGCTGATGTGTCTCTGCTGAGAAACCCCTGGCTGCAGACACAGCC	1483
QY	1203	CTTTACATCTTGATCCGGACATGTGACGTGTGTCTAGAGACGTCCTGTTCCGCTGACAT	1262
Db	1484	CTTTACATCTTGATCCGGACATGTGACGTGTGTCTAGAGACGTCCTGTTCCGCTGACAT	1543
QY	1263	CTTCCAAATGCAAGCCACGACCCGCTAACCTGGGGCCTATTACATTTTCCGANTCAAAATC	1322
Db	1544	CTTCCAAATGCAAGCCACGACCCGCTAACCTGGGGCCTATTACATTTTCCGANTCAAAATC	1603
QY	1323	TGGGAATGAGGGCAGAGAAATTTTACATGCGGCAAAAGGGGCCCATCAGTGCACCCCTGCT	1382
Db	1604	TGGGAATGAGGGCAGAGAAATTTTACATGCGGCAAAAGGGGCCCATCAGTGCACCCCTGCT	1663
QY	1383	GATGACAGCCCCCATCAAAAGGGCCCCGGGAAATCCAGCTGCAGCTTGGAATATCATCTGT	1442
Db	1664	GATGACAGCCCCCATCAAAAGGGCCCCGGGAAATCCAGCTGCAGCTTGGAATATCATCTGT	1723
QY	1443	CACACTGTATCAACTTCAGAGGACGTCCTGATCCGACGCGGAGATATATGTGCGCA	1502
Db	1724	CACACTGTATCAACTTCAGAGGACGTCCTGATCCGACGCGGAGATATATGTGCGCA	1783
QY	1503	GTACCCATTTCTGAGCTTCGGGCTTGAGACCTTCGACGCTGCTCTCATTTGGACCAAGGGA	1562
Db	1784	GTACCCATTTCTGAGCTTCGGGCTTGAGACCTTCGACGCTGCTCTCATTTGGACCAAGGGA	1843
QY	1563	CAGGACAAAGAGAGAAATTAACAGAGAAATAGAGCGACACAGAGGTATGGCATTTCTGT	1622
Db	1844	CAGGACAAAGAGAGAAATTAACAGAGAAATAGAGCGACACAGAGGTATGGCATTTCTGT	1903
QY	1623	CTGAAGCTTTCCCGAAGAGTACAGCCCGACTCTCTGACTCTACCTGTACTATTGCGAGA	1682
Db	1904	CTGAAGCTTTCCCGAAGAGTACAGCCCGACTCTCTGACTCTACCTGTACTATTGCGAGA	1963
QY	1683	CCGTGCACCTGCAGGACTTCCACCCCACTTCTATGATACAGTATCAAAAGATTT	1742
Db	1964	CCGTGCACCTGCAGGACTTCCACCCCACTTCTATGATACAGTATCAAAAGATTT	2023
QY	1743	ATCATTTGCTCCCTGATAGAAATGTTGTAATTTTCAAGGCCCTTCAGTTATTATTCGA	1802
Db	2024	ATCATTTGCTCCCTGATAGAAATGTTGTAATTTTCAAGGCCCTTCAGTTATTATTCGA	2083
QY	1803	CTATTTTCAAGAAATATGATTTAGTTGTTCGGGGGCTCTAGATATGTTCAAGACGTG	1862
Db	2084	CTATTTTCAAGAAATATGATTTAGTTGTTCGGGGGCTCTAGATATGTTCAAGACGTG	2143
QY	1863	AACAGCTTGCTGCTACTCTTCACCTCTTCCACTCCTTCTCAGCTGTGTACTGCTTG	1922
Db	2144	AACAGCTTGCTGCTACTCTTCACCTCTTCCACTCCTTCTCAGCTGTGTACTGCTTG	2203
QY	1923	CAAAAGACCCGGGAGCTGGCGGGGGAACCTGGGAGTAGTATGTTTGCTTTTGGGTACACA	1982
Db	2204	CAAAAGACCCGGGAGCTGGCGGGGGAACCTGGGAGTAGTATGTTTGCTTTTGGGTACACA	2263
QY	1983	GAGAAAGGTATGTAAACAACAACACAGCAGAGATCGAAGGGTTTTAGAGAAATGTCTTCAA	2042
Db	2264	GAGAAAGGTATGTAAACAACAACACAGCAGAGATCGAAGGGTTTTAGAGAAATGTCTTCAA	2323
QY	2043	AACCATGCTGGTATTTTCAACCATAAAGAAGTTTCAGTTGCTTTAAATTTGTATATAC	2102
Db	2324	AACCATGCTGGTATTTTCAACCATAAAGAAGTTTCAGTTGCTTTAAATTTGTATATAC	2383
QY	2103	GGTTTAATTCGTCTTGTTCATTTTGAGATTTTAAAAAATATGTGTGAATTCCTTC	2162
Db	2384	GGTTTAATTCGTCTTGTTCATTTTGAGATTTTAAAAAATATGTGTGAATTCCTTC	2443
QY	2163	GAAGAGCCTTCAGACATGATGTATGTCTTCTCCAAACCCAGTCCGATCCATATTT	2222
Db	2444	GAAGAGCCTTCAGACATGATGTATGTCTTCTCCAAACCCAGTCCGATCCATATTT	2503

OY	2223	AGCCCACTGTTTCTTTTGAGGACCCCTTAATCTGCTTCCTTGAAATTTTACCACATT	2282
Db	2504	AGCCCACTGTTTCTTTTGAGAGACCCCTTAATCTGCTTCCTTGAAATTTTACCACATT	2563
OY	2283	GGATTGGAATGCAGAGGTCTCCAAACTGATTAAATATTGGAAGA	2328
Db	2564	GGATTGGAATGCAGAGGTCTCCAAACTGATTAAATATTGGAAGA	2609
 RESULT 15 US-10-180-557-407			
Sequence 407, Application US/40180557			
Publication No. US20030022301A1			
GENERAL INFORMATION:			
APPLICANT: Baker, Kevin P.			
APPLICANT: Chen, Jian			
APPLICANT: Desnoyers, Luc			
APPLICANT: Goddard, Audrey			
APPLICANT: Godowski, Paul J.			
APPLICANT: Gurney, Austin L.			
APPLICANT: Pap, James			
APPLICANT: Smith, Victoria			
APPLICANT: Watanabe, Colin K.			
APPLICANT: Wood, William I.			
APPLICANT: Zhang, Zemin			
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC			
TITLE OF INVENTION: ACIDS ENCODING THE SAME			
FILE REFERENCE: P3430R1C147			
CURRENT FILING DATE: 2002-06-25			
Prior Application removed - See File Wrapper or Palm			
NUMBER OF SEQ ID NOS: 612			
SEQ ID NO 407			
LENGTH: 2609			
TYPE: DNA			
ORGANISM: Homo Sapien			
US-10-180-557-407			
 Query Match 99.8%; Score 2322.8; DB 9; Length 2609;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 2324; Conservative 0; Mismatches 2; Indels 0; Gaps 0			
OY	3	CCCCGGCGTCGCCGGGTGTCCTCTCCAGACATCGCTCGGCCCTCTGGAATAAACACCC	62
Db	284	CCCCGGCGTCCTCCCGTGTCTCTCCAGACATCGCTCGGCCCTCTGGAATAAACACCC	343
OY	63	GCGAGCCCCGAGGGCCCGACAGAGAGCGCACGTGCCGACTCTCCGGGGGTCCGCCCG	122
Db	344	GCGAGCCCCGAGGGCCCGACAGAGAGCGCACGTGCCGACTCTCCGGGGGTCCGCCCG	403
OY	123	CGAGCTTCTTCGCGCTTCGCAATCTGCATCTGCGCGCTTCGCGCTTCGCAATCTGCAATGAA	182
Db	404	CGAGCTTCTTCGCGCTTCGCAATCTGCATCTGCGCGCTTCGCGCTTCGCAATGCAAGAAATAA	463
OY	183	AAGGATACACTGTTACCATTTGCGCTCTGTGTTCACAAGCCCTGGGAATGCACAGGC	242
Db	464	AAGGATACACTGTTACCATTTGCGCTCTGTGTTCACAAGCCCTGGGAATGCACAGGC	523
OY	243	ACAGTCCAGAAAGGCTTTGACCTGTGATCCGACGTCCAGTAGCACTGTTAGATATTGATGA	302
Db	524	ACAGTCCAGAAAGGCTTTGACCTGTGATCCGACGTCCAGTAGCACTGTTAGATATTGATGA	583
OY	303	ATCCGCAACCATCCCGAGGCGTCGCGAGAGACATGATGTGTTAACCAAATGGCGG	362
Db	584	ATCCGCAACCATCCCGAGGCGTCGCGAGAGACATGATGTGTTAACCAAATGGCGG	643
OY	363	GTAATTATGCAATCCCGGACAAACCTGTGTATCGAGGGCCCTACTCGGAACCCATATC	422
Db	644	GTAATTATGCAATCCCGGACAAACCTGTGTATCGAGGGCCCTACTCGGAACCCATATC	703
OY	423	GACCCCTACTCAAGGTCCGTAACGAGAGGTGCCCCACCACTCTCAGACTCCAACATATCC	482
Db	704	GACCCCTACTCAAGGTCCGTAACGAGAGGTGCCCCACCACTCTCAGACTCCAACATATCC	766

OY	1803	CTATTTTCAAGAAAATAGATTAGTGTGGGGGGGTGAGTCTAFTGTCAAAACGTG	1862
Db	2084	CTATTTTCAAGAAAATAGATTAGTGTGGGGGGGTGAGTCTAFTGTCAAAACGTG	2143
OY	1863	AACAGCTTGCTGTCACTTCCTACCTCTCCACTCTCTCTCACTGTGTACTGCTTG	1922
Db	2144	AACGCTTGTGTCTACCTTCCTACCTCTTCACCTCTCTCTCACTGTGTACTGCTTG	2203
OY	1923	CAAGACCCGGGACCTGGGGGAAACCTGGGATCTAGTGTGCTTTTGCTGTACACA	1982
Db	2204	CAAGACCCGGGACCTGGGGGAAACCTGGGATCTAGTGTGCTTTTGCTGTACACA	2263
OY	1983	GAGAAGCCTATGTATTAACAAACACAGACAGATCGAAGGTTTTTATGAAATGTCTCAA	2042
Db	2264	GAGAAGCCTATGTATTAACAAACACAGACAGATCGAAGGTTTTTATGAAATGTCTCAA	2323
OY	2043	AACCATGCGTGTTTTCACACCAATTAAGAAGTTTCAATGTCCTTAATTTATATAC	2102
Db	2324	AACCATGCGTGTTTTCACACCAATTAAGAAGTTTCAATGTCCTTAATTTATATAC	2383
OY	2103	GGTTTAATTCGTCTGTCTCAATTTTGAATATTTTAAAAAATATGTGTAGTAATTCCTC	2162
Db	2384	GGTTTAATTCGTCTGTCTCAATTTTGAATATTTTAAAAAATATGTGTAGTAATTCCTC	2443
OY	2163	GAAAGGCGTTCAGACACATGCTATGTCGTCTCCAAACCCAGTCCCTGCTCCATTTT	2222
Db	2444	GAAAGGCGTTCAGACACATGCTATGTCGTCTCCAAACCCAGTCCCTGCTCCATTTT	2503
OY	2223	AGCCCACTGTTTCTTTGAGAGACCCCTTAATCTTGCTTTCTTTTGAATTTTACCCAAT	2282
Db	2504	AGCCCACTGTTTCTTTGAGAGACCCCTTAATCTTGCTTTCTTTTGAATTTTACCCAAT	2563
OY	2283	GGATTGGAATCGAAGGCTTCCAAACGATTAATATTTTGAAGGA 2328	
Db	2564	GGATTGGAATCGAAGGCTTCCAAACGATTAATATTTTGAAGGA 2609	

APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C66
CURRENT FILING DATE: 2002-06-20
Prior Application removed - see file Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 407
LENGTH: 2609
TYPE: DNA
ORGANISM: Homo Sapien
US-10-176-913-407

Query Match 99.8%; Score 2322.8; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2324; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

3 CCGGGGCTCTCCCGCTGCTCTCCACGACGCTGCGCCCTCTGGAATAAACACCC 62
284 CCGGGGCTCTCCCGCTGCTCTCCACGACGCTGCGCCCTCTGGAATAAACACCC 343
63 GCGAGCCCGAGGCGCCAGAGAGGCGGAGCTGCGCCGAGCTCTCCGGGGTCCCGCCG 122
344 GCGAGCCCGAGGCGCCAGAGAGGCGGAGCTGCGCCGAGCTCTCCGGGGTCCCGCCG 403
123 CGAGCTTTTCTTCCGCTTTCGATCTCTCTCTGCGCGCTCTTGGACATGCCAGAAATAA 182
404 CGAGCTTTTCTTCCGCTTTCGATCTCTCTCTGCGCGCTCTTGGACATGCCAGAAATAA 463
183 AAGGATCTACATGTTACATCTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 242
464 AAGGATCTACATGTTACATCTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 523
243 ACAGTGCAGCAATGAGCTTTGACCTGATGCGCAGTCAAGACATGTTAGATATTGATGA 302
524 ACAGTGCAGCAATGAGCTTTGACCTGATGCGCAGTCAAGACATGTTAGATATTGATGA 583
303 ATGCCGACCAATCCCGAGGCTTCCGAGAGACATGATGTTAAACCAAAATGGCGG 362
584 ATGCCGACCAATCCCGAGGCTTCCGAGAGACATGATGTTAAACCAAAATGGCGG 643
363 GTATTATGATTCCTCCCGAGCAACCTGTGTATCGAGGGCCCTACTGGAACCCCTACTC 422
644 GTATTATGATTCCTCCCGAGCAACCTGTGTATCGAGGGCCCTACTGGAACCCCTACTC 703
423 GACCCCTACTAGTCCGATACCCAGAGCTGCCCCACACTCTCAGTCTCAAACTATCC 482
704 GACCCCTACTAGTCCGATACCCAGAGCTGCCCCACACTCTCAGTCTCAAACTATCC 763
483 CACGATCTCCAGGCTCTTATATGCGCTTTGGATACAGATGATGAAGCAACCAATG 542
764 CACGATCTCCAGGCTCTTATATGCGCTTTGGATACAGATGATGAAGCAACCAATG 823
543 TGTGGATGTGAGAGATGTGCAACAGATTCGCAACAGTGCACCAACCCAGATGTGAT 602
824 TGTGGATGTGAGAGATGTGCAACAGATTCGCAACAGTGCACCAACCCAGATGTGAT 883
603 CAATATGAAGGGGTACACCTGCTCTCTCAGCAGAGATGATGCTTCTGAAGGCA 662
884 CATATCTGAAGGGGTACACCTGCTCTCTCAGCAGAGATGATGCTTCTGAAGGCA 943
663 GTCTTAGACATGATGATGTGCTATGTGTTACTGCGACAGCTCTGTGGAATGTTCC 722
|||||

944 GTGCTTAGACATGATGATGTGCTATGTTACTGCCAGACCTCTGTGGAATGTTCC 1003
723 TGGATCTATTTCTGATCATGCAACCTGGTTTTACCTCATGAGATGAGAGCTCTTG 782
1004 TGGATCTATTTCTGATCATGCAACCTGGTTTTACCTCATGAGATGAGAGCTCTTG 1063
783 CCAAGATGTGAACAGATGTGCCACCGAAGACCCCTGGCTGCAAACTGCTGACACCTA 842
1064 CCAAGATGTGAACAGATGTGCCACCGAAGACCCCTGGCTGCAAACTGCTGACACCTA 1123
843 CGGCTCTTATATGCGCGCTGACCCAGATATGATTAAGTAAAGATGGCGTCAATG 902
1124 CGGCTCTTATATGCGCGCTGACCCAGATATGATTAAGTAAAGATGGCGTCAATG 1183
903 CAGTGAATGAGAGAGTGCAGCTCTCTGATGTTCTCTGCAACATGATGTGTGAACA 962
1184 CAGTGAATGAGAGAGTGCAGCTCTCTGATGTTCTCTGCAACATGATGTGTGAACA 1243
963 GCCCGACATCTTCTGCTCTGCTCCCTCCAGGCTACATCTCTGTGATGACCAACGAAG 1022
1244 GCCCGACATCTTCTGCTCTGCTCCCTCCAGGCTACATCTCTGTGATGACCAACGAAG 1303
1023 CTGCCAAGACATCAACGATGTGACACAGGACACACGTCGACACCTGACACACGTCG 1082
1304 CTGCCAAGACATCAACGATGTGACACAGGACACACGTCGACACCTGACACACGTCG 1363
1083 CTACAAATTTCAAGGGGGCTTCAAAATGATGATGACCCCTGCGCTGAGAGACCTTATCT 1142
1364 CTACAAATTTCAAGGGGGCTTCAAAATGATGATGACCCCTGCGCTGAGAGACCTTATCT 1423
1143 GAGATAGTATGATTAACCGCTGATGTCCTGCTGAGAAACCTGCTGACAGACACGCT 1202
1424 GAGATAGTATGATTAACCGCTGATGTCCTGCTGAGAAACCTGCTGACAGACACGCT 1483
1203 CTTTACCATTTGATACCGGGACATGAGAGTGTGTGACAGAGAGTCTCCGCTGACAT 1262
1484 CTTTACCATTTGATACCGGGACATGAGAGTGTGTGACAGAGAGTCTCCGCTGACAT 1543
1263 CTTTCAAAATGCAAGCCAGACCCGCTACCTGCGGGCTTATTACATTTTCCAGATCAATC 1322
1544 CTTTCAAAATGCAAGCCAGACCCGCTACCTGCGGGCTTATTACATTTTCCAGATCAATC 1603
1323 TGGGAATGAGGCGAGAGATTTTACATGCGGCAAAAGGGCCCATCACTGACACCTGCT 1382
1604 TGGGAATGAGGCGAGAGATTTTACATGCGGCAAAAGGGCCCATCACTGACACCTGCT 1663
1383 GATGACACCGCCCATCAAGGGGCCCGGGAATCAGTGTGACTGGAATGATCACTGT 1442
1664 GATGACACCGCCCATCAAGGGGCCCGGGAATCAGTGTGACTGGAATGATCACTGT 1723
1443 CAACACTGTATCAACTTCAAGAGCAGCTCCGTCATCCGATGCGGATATATGTGCGCA 1502
1724 CAACACTGTATCAACTTCAAGAGCAGCTCCGTCATCCGATGCGGATATATGTGCGCA 1783
1503 GTACCATTTGAGCTCTGGGCTGGAAGCTCTCCAGAGCTGCTCTGATTTGGACCAAGGA 1562
1784 GTACCATTTGAGCTCTGGGCTGGAAGCTCTCCAGAGCTGCTCTGATTTGGACCAAGGA 1843
1563 CAGGAGAAGAGAGAAATTAACAGAGAGAAATGAGCGGACACAGAGCTTAAGCATTTCCG 1622
1844 CAGGAGAAGAGAGAAATTAACAGAGAGAAATGAGCGGACACAGAGCTTAAGCATTTCCG 1903
1623 CTGAAGCTTTCCCGAAGAGTGCAGCCCGGACTCTCTGACTCTGACCTGACTATTGACGA 1682
1904 CTGAAGCTTTCCCGAAGAGTGCAGCCCGGACTCTCTGACTCTGACCTGACTATTGACGA 1963
1683 CCTGTACCTCTGAGAGCTTGCACACCCCACTCTCTATGATACATTAACAAAAGTAT 1742
1964 CCTGTACCTCTGAGAGCTTGCACACCCCACTCTCTATGATACATTAACAAAAGTAT 2023
1743 ATCATTTCTCTCTGATGAGAGATTTGTTGGATTTTCAAGGCTTCAAGTTATTTCGA 1802
2024 ATCATTTCTCTCTGATGAGAGATTTGTTGGATTTTCAAGGCTTCAAGTTATTTCGA 2083
|||||

```

Db 524 ACAGTGCAGAAATGGCTTTGACCTGATCCGACGTGACGAGACGTGTTAGATATGATGA 583
Qy 303 ATGCCGAACCATCCCGGAGGCTCCGAGAGACATATGTGTATTAAACAAATGGGG 362
Db 584 ATGCCGAACCATCCCGGAGGCTCCGAGAGACATATGTGTATTAAACAAATGGGG 643
Qy 363 GTATTTATGATTCCTCCGAGCAACCTGTGTATCGAGGGGCTACTGAAACCCCTACTGC 422
Db 644 GTATTTATGATTCCTCCGAGCAACCTGTGTATCGAGGGGCTACTGAAACCCCTACTGC 703
Qy 423 GACCCCTCACTCAGATCCGATACCCAGAGCTGCCCCACACACTCAAGCTCCAAATATCC 482
Db 704 GACCCCTCACTCAGATCCGATACCCAGAGCTGCCCCACACACTCAAGCTCCAAATATCC 763
Qy 483 CACGATCTCCAGGCTCTTATATGCGGCTTTGGATACCAGATGATGAAGCAACCAATG 542
Db 764 CACGATCTCCAGGCTCTTATATGCGGCTTTGGATACCAGATGATGAAGCAACCAATG 823
Qy 543 TGTGATGTGAGCAGATGTGCAACAGATCCACACAGTGCACCCACCCAGATCTGCAT 602
Db 824 TGTGATGTGAGCAGATGTGCAACAGATCCACACAGTGCACCCACCCAGATCTGCAT 883
Qy 603 CAATCTGAAGCGGGGTACACCTGCTCTGACCCGAGGATATTTGGCTTGTGAAGGCCA 662
Db 884 CAATCTGAAGCGGGGTACACCTGCTCTGACCCGAGGATATTTGGCTTGTGAAGGCCA 943
Qy 663 GTGCTAGACATTTGATGAATGTGCTATGTTACTGACAGAGCTGTGGGAATGTTC 722
Db 944 GTGCTAGACATTTGATGAATGTGCTATGTTACTGACAGAGCTGTGGGAATGTTC 1003
Qy 723 TGGATCTTATTTCTGTACATGCAACCTGTGTTTACCCTCAATGAGATGGAAGGCTTGG 782
Db 1004 TGGATCTTATTTCTGTACATGCAACCTGTGTTTACCCTCAATGAGATGGAAGGCTTGG 1063
Qy 783 CCAAGATGTGAGAGATGTGCTCCACCGGAACCCCTGGGTGCAAACTGCTGCAACCTGA 842
Db 1064 CCAAGATGTGAGAGATGTGCTCCACCGGAACCCCTGGGTGCAAACTGCTGCAACCTGA 1123
Qy 843 CGGCTCTTCATCTGCGCTGTGACCCAGAGATGATGAAGTGAAGATGGGCTTCATTTG 902
Db 1124 CGGCTCTTCATCTGCGCTGTGACCCAGAGATGATGAAGTGAAGATGGGCTTCATTTG 1183
Qy 903 CAGTGATATGAGCAGATGTGACGCTTCTGTGAGTTCTCTGCAACATGAGTGTGAACCA 962
Db 1184 CAGTGATATGAGCAGATGTGACGCTTCTGTGAGTTCTCTGCAACATGAGTGTGAACCA 1243
Qy 963 GCGCCGACATATCTTCTGCTCCCTCCAGGCTACATCTGCTGTGATGACACCGAAG 1022
Db 1244 GCGCCGACATATCTTCTGCTCCCTCCAGGCTACATCTGCTGTGATGACACCGAAG 1303
Qy 1023 CTGCCAAGACATCAAGATGTGAGCAGAGAACACAGCTGCAACCTGCAAGCAGACAGTG 1082
Db 1304 CTGCCAAGACATCAAGATGTGAGCAGAGAACACAGCTGCAACCTGCAAGCAGACAGTG 1363
Qy 1083 CTACAAATTTACAAGAGGGGCTTCAAAATGCATGACCCCATCCGCTGTGAAGAGCTTATCT 1142
Db 1364 CTACAAATTTACAAGAGGGGCTTCAAAATGCATGACCCCATCCGCTGTGAAGAGCTTATCT 1423
Qy 1143 GAGGATCAGTATACCGCTGTATGTCTGCTGTGAAGAACCTGGCTGCAAGACAGCC 1202
Db 1424 GAGGATCAGTATACCGCTGTATGTCTGCTGTGAAGAACCTGGCTGCAAGACAGCC 1483
Qy 1203 CTTTACATCTTGTACCGGACATGAGAGTGTGTCAGAGAGCTCCGTTCCGGCTGAGAT 1262
Db 1484 CTTTACATCTTGTACCGGACATGAGAGTGTGTCAGAGAGCTCCGTTCCGGCTGAGAT 1543
Qy 1263 CTTCGAATGCAAGCAGACGACCGCTACCTGGGGCTTATTAATTTCCAGATCAATC 1322
Db 1544 CTTCGAATGCAAGCAGACGACCGCTACCTGGGGCTTATTAATTTCCAGATCAATC 1603
Qy 1323 TGGGAATGAGGCGAGAGATTTTACATGCGCAACGCGCCCATCACTGCCACCTGGT 1382

```

```

Db 1604 TGGGAATGAGGCGAGAGATTTTACATGCGCAACGCGCCCATCACTGCCACCTGGT 1663
Qy 1383 GATGACAGCGCCCATCAAGAGGCGCCCGGAATCCAGCTGAGCTGSAATGATCACTGT 1442
Db 1664 GATGACAGCGCCCATCAAGAGGCGCCCGGAATCCAGCTGAGCTGSAATGATCACTGT 1723
Qy 1443 CAACACTGTATCACTTCAAGAGGAGAGCTCCGTGTATCCGAGCTGCGATATATGTGCGCA 1502
Db 1724 CAACACTGTATCACTTCAAGAGGAGAGCTCCGTGTATCCGAGCTGCGATATATGTGCGCA 1783
Qy 1503 GTACCCATTTCTGACCTTCGGGCTGAGGCTCCGACGTGCTCTCATTTGGACCAAGGA 1562
Db 1784 GTACCCATTTCTGACCTTCGGGCTGAGGCTCCGACGTGCTCTCATTTGGACCAAGGA 1843
Qy 1563 CAGAGAGAGAGGAATTAACAGAGAGATAGAGCGACACAGAGTATGAGATTTCCG 1622
Db 1844 CAGAGAGAGAGGAATTAACAGAGAGATAGAGCGACACAGAGTATGAGATTTCCG 1903
Qy 1623 CTGAACGTTTCCCGGAGAGAGTCAAGCCCGGACTTCTGACTCTCACCTGACTATTCGAGA 1682
Db 1904 CTGAACGTTTCCCGGAGAGAGTCAAGCCCGGACTTCTGACTCTCACCTGACTATTCGAGA 1963
Qy 1683 CCTGTACCTTCGACGAGCTTCCACCCCACTTCTATGATATCACTTATCAAAAGTAT 1742
Db 1964 CCTGTACCTTCGACGAGCTTCCACCCCACTTCTATGATATCAAAAGTAT 2023
Qy 1743 ATCATGCTCCCGCTGATAGAGATTTGTTGATGAATTTTCAAGGCTTCAGTTATTTCCA 1802
Db 2024 ATCATGCTCCCGCTGATAGAGATTTGTTGATGAATTTTCAAGGCTTCAGTTATTTCCA 2083
Qy 1803 CTATTTTCAAGAAATATGATTTAGTTTGGGGGCTGAGTCTATGTTCAAAAGCTGTG 1862
Db 2084 CTATTTTCAAGAAATATGATTTAGTTTGGGGGCTGAGTCTATGTTCAAAAGCTGTG 2143
Qy 1863 AAGAGCTGTGACCTCTTCACTTCTTCCACCTCTCTGACAGTGTACTGTGTTG 1922
Db 2144 AAGAGCTGTGACCTCTTCACTTCTTCCACCTCTCTGACAGTGTACTGTGTTG 2203
Qy 1923 CAAAGACCCGAGAGCTGGCGGGGAACCTGGGAGTAGTATGTTCTTTTCCGTACAGA 1982
Db 2204 CAAAGACCCGAGAGCTGGCGGGGAACCTGGGAGTAGTATGTTCTTTTCCGTACAGA 2263
Qy 1983 GAGAGGCTATGTAAACAAACACAGAGATCGAAGGTTTGAAGATGTGTTCAA 2042
Db 2264 GAGAGGCTATGTAAACAAACACAGAGATCGAAGGTTTGAAGATGTGTTCAA 2323
Qy 2043 AACCATGCTGTGTTTCAACCAATTAAGAAAGTTCAAGTTCCCTTAATTTGATTAAC 2102
Db 2324 AACCATGCTGTGTTTCAACCAATTAAGAAAGTTCAAGTTCCCTTAATTTGATTAAC 2383
Qy 2103 GGTATTAATCTGTGTTTCAATTTTGAAGTATTTTAAATAATATGCTGAGATTCCTTC 2162
Db 2384 GGTATTAATCTGTGTTTCAATTTTGAAGTATTTTAAATAATATGCTGAGATTCCTTC 2443
Qy 2163 GAAAGGCTTCAGACATGCTATGTTCTGTTTCCAAACCCAGCTCTCTCCATTTT 2222
Db 2444 GAAAGGCTTCAGACATGCTATGTTCTGTTTCCAAACCCAGCTCTCTCCATTTT 2503
Qy 2223 AGCCCATGTTTCTTTGAGAGACCCCTTAATCTGCTTCTTAAATTTTAAACCAAT 2282
Db 2504 AGCCCATGTTTCTTTGAGAGACCCCTTAATCTGCTTCTTAAATTTTAAACCAAT 2563
Qy 2283 GGAATGGAATGACAGGCTTCCAAAGTATTAATATTTGAAGAGA 2328
Db 2564 GGAATGGAATGACAGGCTTCCAAAGTATTAATATTTGAAGAGA 2609

```

RESULT 13
 US-10-176-913-407
 ; Sequence 407, Application US/10176913
 ; Publication No. US20030022298A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.

|||||
Db 2504 AGCCACTGTTCTTTGAGAGACCCTTAATCTTCTCTTAAGATTTTACCATT 2563
OY 2283 GGATTGGAATGCAGAGCTCTCCAAACTATTAATTTTGAAGA 2328
Db 2564 GGATTGGAATGCAGAGCTCTCCAAACTATTAATTTTGAAGA 2609

RESULT 10
US-10-175-752-407
; Sequence 407, Application US/10175752
; Publication No. US20030022295A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Guiney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Matanabe, Collin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zhenli
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C60
; CURRENT APPLICATION NUMBER: US/10/175,752
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 407
; LENGTH: 2609
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-175-752-407

Query Match 99.8%; Score 2322.8; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2324; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 CCCGGGCTCTCCCGCTGCTCTCCAGAGCTGCTGGGCCCCCTGTGAATAAACACCC 62
Db 284 CCCGGGCTCTCCCGCTGCTCTCCAGAGCTGCTGGGCCCCCTGTGAATAAACACCC 343
OY 63 GCGAGCCCGAGAGGCCAGAGAGAGCCGAGCTCTCCGGGGGCTCCGCGCG 122
Db 344 GCGAGCCCGAGAGGCCAGAGAGAGCCGAGCTCTCCGGGGGCTCCGCGCG 403
OY 123 CGAGCTTTTCTCGGCTTGCATCTCTGCGCGCTTGTGACATGCCAGGAATAA 182
Db 404 CGAGCTTTTCTCGGCTTGCATCTCTGCGCGCTTGTGACATGCCAGGAATAA 463
OY 183 AAGGATCTACATGTTACCATTTGGCTCTGTCTTCCAAAGCCCTGGGAATGACAGGC 242
Db 464 AAGGATCTACATGTTACCATTTGGCTCTGTCTTCCAAAGCCCTGGGAATGACAGGC 523
OY 243 ACAGTGCAGGAATGGCTTTGACCTGATGCCAGTCCAGTACAGAGTGTAGATTTAGTA 302
Db 524 ACAGTGCAGGAATGGCTTTGACCTGATGCCAGTCCAGTACAGAGTGTAGATTTAGTA 583
OY 303 ATGCCGAACATCCCGAGAGCGCTGCCGAGAGACATGATGTGTTAACCAAAATGGCGG 362
Db 584 ATGCCGAACATCCCGAGAGCGCTGCCGAGAGACATGATGTGTTAACCAAAATGGCGG 643
OY 363 GATTATTATGATTTCCCGGAGCAAAACCTGTATGAGAGGCCCTACTCGAAACCCCTACTC 422
Db 644 GATTATTATGATTTCCCGGAGCAAAACCTGTATGAGAGGCCCTACTCGAAACCCCTACTC 703
OY 423 GACCCCTACTAGTCCGTACCCAGACAGTGGCCCACTCTACAGTCCAAACTATTC 482
Db 704 GACCCCTACTAGTCCGTACCCAGACAGTGGCCCACTCTACAGTCCAAACTATTC 763

OY 483 CAGATCTCCAGGCTCTTATATGCGCTTTGGATACAGATGATGAAGCAACCAATG 542
Db 764 CAGATCTCCAGGCTCTTATATGCGCTTTGGATACAGATGATGAAGCAACCAATG 823
OY 543 TGTGATGTGAGAGAGTGTGACACAGATTTCCACAGTGCACAAACCCACAGATGTGAT 602
Db 824 TGTGATGTGAGAGAGTGTGACACAGATTTCCACAGTGCACAAACCCACAGATGTGAT 883
OY 603 CAATPACTGAAGGGGGTACACCTGCTCTGCACCGAGAGATTTGGCTTGTGAAGGCA 662
Db 884 CAATPACTGAAGGGGGTACACCTGCTCTGCACCGAGAGATTTGGCTTGTGAAGGCA 943
OY 663 GTGCTTAGACATGATGAATGTGCTATGTTACTGTCAGAGAGCTCTGTGGAAATGTTCC 722
Db 944 GTGCTTAGACATGATGAATGTGCTATGTTACTGTCAGAGAGCTCTGTGGAAATGTTCC 1003
OY 723 TGGATCCATTCTTGTATACATGCAACCTGTTTACCCTCATGAGATGAGTGAAGGCTTGG 782
Db 1004 TGGATCCATTCTTGTATACATGCAACCTGTTTACCCTCATGAGATGAGTGAAGGCTTGG 1063
OY 783 CCAAGATGTGAAGAGTGTGCGCACCGAGAAACCCCTGCGTGAACCTGCTCAACACCTA 842
Db 1064 CCAAGATGTGAAGAGTGTGCGCACCGAGAAACCCCTGCGTGAACCTGCTCAACACCTA 1123
OY 843 CGGCTCTTTCATGTCGCGGCTGTGACCCAGATATGAGGAAATGGCGTTCATTTG 902
Db 1124 CGGCTCTTTCATGTCGCGGCTGTGACCCAGATATGAGGAAATGGCGTTCATTTG 1183
OY 903 CAGTGTATGAGAGAGTGCAGCTTCTGTGATTTCCCTGTCACATGATGATGAGAACCA 962
Db 1184 CAGTGTATGAGAGAGTGCAGCTTCTGTGATTTCCCTGTCACATGATGATGAGAACCA 1243
OY 963 GCGCGGACATATTTGCTGCTCTGCTCCAGAGTATATCTGCTGATGACAAACCGAG 1022
Db 1244 GCGCGGACATATTTGCTGCTCTGCTCCAGAGTATATCTGCTGATGACAAACCGAG 1303
OY 1023 CTGCAAGACATCAAGATGTGAGCAGAGAAACACAGGTCACCTGCGAGAGAGTGTG 1082
Db 1304 CTGCAAGACATCAAGATGTGAGCAGAGAAACACAGGTCACCTGCGAGAGAGTGTG 1363
OY 1083 CTACAAATTTCAAGGGGGCTTCAAAATGATGACACCCATCGCTGTGAGAGAGCTTATCT 1142
Db 1364 CTACAAATTTCAAGGGGGCTTCAAAATGATGACACCCATCGCTGTGAGAGAGCTTATCT 1423
OY 1143 GAGGATCAGTGAATACCGGCTGTATGTCCTGCTGANAACCTGCTGACAGACAGCC 1202
Db 1424 GAGGATCAGTGAATACCGGCTGTATGTCCTGCTGANAACCTGCTGACAGACAGCC 1483
OY 1203 CTTTACCATCTTGTACCGGACATGAGACGTGTGTGAGAGAGCTCCGTCCTGACAT 1262
Db 1484 CTTTACCATCTTGTACCGGACATGAGACGTGTGTGAGAGAGCTCCGTCCTGACAT 1543
OY 1263 CTTTCAAAATGCAAGCCGACAGCCGCTACCTGCGGCTATTACATTTTCCAGATCAATC 1322
Db 1544 CTTTCAAAATGCAAGCCGACAGCCGCTACCTGCGGCTATTACATTTTCCAGATCAATC 1603
OY 1323 TGGATGTAGGGGAGAGATTTTACATGGGGGCAAAAGGGGCCCATCATGATGCCACCCGT 1382
Db 1604 TGGATGTAGGGGAGAGATTTTACATGGGGGCAAAAGGGGCCCATCATGATGCCACCCGT 1663
OY 1383 GATGACAGCCGCCATCAAAAGGGGCCCGGGAATTCACAGCTGAGTGTGAATGATCACTGT 1442
Db 1664 GATGACAGCCGCCATCAAAAGGGGCCCGGGAATTCACAGCTGAGTGTGAATGATCACTGT 1723
OY 1443 CAACACTGTATCAACTTGAAGGACAGCTGCTGATCCGACTGCGGATATATGTGCGCA 1502
Db 1724 CAACACTGTATCAACTTGAAGGACAGCTGCTGATCCGACTGCGGATATATGTGCGCA 1783
OY 1503 GTACCATTTGAGGCTGCGGGGAGGAGCCGAGAGCTGCTCATTTGGCACCAGAGGA 1562
Db 1784 GTACCATTTGAGGCTGCGGGGAGGAGCCGAGAGCTGCTCATTTGGCACCAGAGGA 1843
OY 1563 CAGGAGAGAGAGGAATTAACAGAGAGATGAGAGAGACACAGACGTTAGGCAATTCTTG 1622

OY	63	GGGAGCCCCGAGGGGCCAGAGGAGGGGGGAGCGTGGCCCGAGCTCTCCGGGGGGTCCGGCCCG	122
Db	344	GGAGGCCCCGAGGGGCCAGAGGAGGGGGAGCGTGGCCCGAGCTCTCCGGGGGGTCCGGCCCG	403
OY	123	CGAGCTTCTTCGCGCTTTCGCAATCTCTCTCCGCGCGCTCTTGAGCATGCCAGGATTA	182
Db	404	CGAGCTTCTTCGCGCTTTCGCAATCTCTCTCCGCGCGCTCTTGAGCATGCCAGGATTA	463
OY	183	AAGGATCTCACTGTTACCAATTTCTGCGTCTCTGTCTCCAAAGCCCTGGGAATGCACAGC	242
Db	464	AAGGATCTCACTGTTACCAATTTCTGCGTCTCTGTCTCCAAAGCCCTGGGAATGCACAGC	523
OY	243	ACAGTGCAGGAATGGCTTTGACCGTGGAAATGCCAGTGCAGAGCAGTGTGATTTATGTA	302
Db	524	ACAGTGCAGGAATGGCTTTGACCGTGGAAATGCCAGTGCAGAGCAGTGTGATTTATGTA	583
OY	303	ATGCCGAACCATCCCGAGGCGCTGCCAGAGAGACATGATGTTGTATTAACCAAAATGGCG	362
Db	584	ATGCCGAACCATCCCGAGGCGCTGCCAGAGAGACATGATGTTGTATTAACCAAAATGGCG	643
OY	363	GTAATTTATGCATTTCCCGGACAAACCTGTGTATTCGAGGGGCCCTACTGAAACCTTACTC	422
Db	644	GTAATTTATGCATTTCCCGGACAAACCTGTGTATTCGAGGGGCCCTACTGAAACCTTACTC	703
OY	423	GACCCCCCTCAAGTGTCCGTAACCCAGACAGTGGCCCCACACCTCTAGCTCCAAACATATCC	482
Db	704	GACCCCCCTCAAGTGTCCGTAACCCAGACAGTGGCCCCACACCTCTAGCTCCAAACATATCC	763
OY	483	CAGATCTCCAGAGCCTCTTATATGTCGCTTGATGATACAGATGAGATGAAGAACAACATG	542
Db	764	CAGATCTCCAGAGCCTCTTATATGTCGCTTGATGATACAGATGAGATGAAGAACAACATG	823
OY	543	TGTGATGTGAGAGAGTGTGCACACAGATTTCCACAGTGCAAACCCACCCAGATCTGCAT	602
Db	824	TGTGATGTGAGAGAGTGTGCACACAGATTTCCACAGTGCAAACCCACCCAGATCTGCAT	883
OY	603	CAATACTGAAGGGGGGTACACCTCTCTCTGACAGCAGGATATTTGGCTTCTGGAAGGCCA	662
Db	884	CAATACTGAAGGGGGGTACACCTCTCTCTGACAGCAGGATATTTGGCTTCTGGAAGGCCA	943
OY	663	GTGCTTAGACATTGATGAATGTGCTATGTTACTGCCAGCAGCTCTGTGCGAATGTTCC	722
Db	944	GTGCTTAGACATTGATGAATGTGCTATGTTACTGCCAGCAGCTCTGTGCGAATGTTCC	1003
OY	723	TGGATCTCTATCTTGTACATGCAACCCGTGTTTACCCCTCATGAGAGTGGAAAGTCTTG	782
Db	1004	TGGATCTCTATCTTGTACATGCAACCCGTGTTTACCCCTCATGAGAGTGGAAAGTCTTG	1063
OY	783	CCAAAGATGCAAGAGAGTGTGCACCGAGAACCCCTGCGTGAACCACTGCGTCAACACTTA	842
Db	1064	CCAAAGATGCAAGAGAGTGTGCACCGAGAACCCCTGCGTGAACCACTGCGTCAACACTTA	1123
OY	843	CGGCTCTTCATCTGCGCGTGTGAACCCAGATATGAACCTTGAAGAAATGGCGTTCATG	902
Db	1124	CGGCTCTTCATCTGCGCGTGTGAACCCAGATATGAACCTTGAAGAAATGGCGTTCATG	1183
OY	903	CAGTGTATGAGAGAGTGCAGCTTCTCTGAAGTCTCTGTGCCAATGATGATGTGTGAACCA	962
Db	1184	CAGTGTATGAGAGAGTGCAGCTTCTCTGAAGTCTCTGTGCCAATGATGATGTGTGAACCA	1243
OY	963	GGCCGGCACATCTTCTGCTCTGCGCCCTCCAGAGTCAATCTGCTGGATGACAAACGAG	1022
Db	1244	GGCCGGCACATCTTCTGCTCTGCGCCCTCCAGAGTCAATCTGCTGGATGACAAACGAG	1303
OY	1023	CTGCGCAAGACATCAACGAATGTGACACAGAACCAACAGTGCACAACTGTGAGCAGAGCTG	1082
Db	1304	CTGCGCAAGACATCAACGAATGTGACACAGAACCAACAGTGCACAACTGTGAGCAGAGCTG	1363
OY	1083	CTACAAATTTCAAGGGGGCTTCAAAATGCATGACCCCATCCGCTGTGAGAGAGCTTATCT	1142
Db	1364	CTACAAATTTCAAGGGGGCTTCAAAATGCATGACCCCATCCGCTGTGAGAGAGCTTATCT	1423

QY	1145	GAGGATCAGGATAAACCGCGTATAGTGCCTGCTAGAAACCCCTGGCTGCAGAACACGCC	1202
Db	1424	GAGGATCAGGATAAACCGCGTATAGTGCCTGCTAGAAACCCCTGGCTGCAGAACACGCC	1483
QY	1203	CTTTACCATCTTGTACCGGGGACATGAGCGTGGTGTCAAGACGCTCCGTTCCGCGTGCAT	1262
Db	1484	CTTTACCATCTTGTACCGGGGACATGAGCGTGGTGTCAAGACGCTCCGTTCCGCGTGCAT	1543
QY	1263	CTTCCAAATCGAAGCCAGACCCGCTACCCCTGGGGCTATTACATTTTCCAGATCCAAATC	1322
Db	1544	CTTCCAAATCGAAGCCAGACCCGCTACCCCTGGGGCTATTACATTTTCCAGATCCAAATC	1603
QY	1323	TGGGAATGAGGGCAGAGAAATTTTACATCGCGCAAAAGGGCCCCATCAGTGCACACCGGT	1382
Db	1604	TGGGAATGAGGGCAGAGAAATTTTACATCGCGCAAAAGGGCCCCATCAGTGCACACCGGT	1663
QY	1383	GATGACAGCCCCATCAAAAGGGCCCCGGAAAAATCCAGCTGAGCTTGGAAATGATCACTGT	1442
Db	1664	GATGACAGCCCCATCAAAAGGGCCCCGGAAAAATCCAGCTGAGCTTGGAAATGATCACTGT	1723
QY	1443	CAACACGTGCATCACTTCAAGAGGAGCTCCGCTGATTCGATCGAGCCGATATATGTGTCGA	1502
Db	1724	CAACACGTGCATCACTTCAAGAGGAGCTCCGCTGATTCGATCGAGCCGATATATGTGTCGA	1783
QY	1503	GTACCCATTTGAGCTTCGGGCTGGAGGCTCCGAGGCTCCGTCTCATTTGGCACCAAGGA	1562
Db	1784	GTACCCATTTGAGCTTCGGGCTGGAGGCTCCGAGGCTCCGTCTCATTTGGCACCAAGGA	1843
QY	1563	CAGGACAGAGAGGAAATTAACAGAGAAATGAGAGCAGACAGACAGCTTAGGCATTTCTGT	1622
Db	1844	CAGGACAGAGAGGAAATTAACAGAGAAATGAGAGCAGACAGACAGCTTAGGCATTTCTGT	1903
QY	1623	CTGAACGTTTCCCGGAAGAGTACAGCCCGGACTCTCGACTCTCACCTGTACTATTCGAGA	1682
Db	1904	CTGAACGTTTCCCGGAAGAGTACAGCCCGGACTCTCGACTCTCACCTGTACTATTCGAGA	1963
QY	1683	CGTGACCCCTGCAGAGCTTGCACCCCGAGTTCTATGATACAGTTATCAAAAAGTATT	1742
Db	1964	CGTGACCCCTGCAGAGCTTGCACCCCGAGTTCTATGATACAGTTATCAAAAAGTATT	2023
QY	1743	ATCATTCCTCCCGATFAGAAATGTTGGTAAATTTTAAAGCCCTYCACTTATTTTCCA	1802
Db	2024	ATCATTCCTCCCGATFAGAAATGTTGGTAAATTTTAAAGCCCTYCACTTATTTTCCA	2083
QY	1803	CTATTTTCAAAAGAAATATGATTAGGTTTGGGGGCTGTAGTCTATGTTCAAGACGTGTG	1862
Db	2084	CTATTTTCAAAAGAAATATGATTAGGTTTGGGGGCTGTAGTCTATGTTCAAGACGTGTG	2143
QY	1863	AACAGCTGCGTCACTCTTCACTCTTCCACTCTTCCACTCTTCTCTACTGTGTACTGCTTGG	1922
Db	2144	AACAGCTGCGTCACTCTTCACTCTTCCACTCTTCCACTCTTCTCTACTGTGTACTGCTTGG	2203
QY	1923	CAAAAGACCCGGAGACTGGCGGGGAACCTGTGGAGTAGCTAGTTTGTCTTTTGGCTACACA	1982
Db	2204	CAAAAGACCCGGAGACTGGCGGGGAACCTGTGGAGTAGCTAGTTTGTCTTTTGGCTACACA	2263
QY	1983	GAGAGGCTATGTAAACAAACACGACGAGATCGAAGGGTTTTAAGATATGTTTCCA	2042
Db	2264	GAGAGGCTATGTAAACAAACACGACGAGATCGAAGGGTTTTAAGATATGTTTCCA	2323
QY	2043	AACCATCGCGTATTTTCAACACATFAAAAGAGTTTCACTGTCTCTTAAATTTGTATTAAC	2102
Db	2324	AACCATCGCGTATTTTCAACACATFAAAAGAGTTTCACTGTCTCTTAAATTTGTATTAAC	2383
QY	2103	GGTTTAATCTGTCTGTCTTCAATTTTGTAGTATTTTAAAAAATATGTCGTAGAAATCCCTTC	2162
Db	2384	GGTTTAATCTGTCTGTCTTCAATTTTGTAGTATTTTAAAAAATATGTCGTAGAAATCCCTTC	2443
QY	2163	GAAAGGCTTCAGACACATGCTATGTTCTGCTTCCCAACCCAGTCTCCTCTCCATTTT	2222
Db	2444	GAAAGGCTTCAGACACATGCTATGTTCTGCTTCCCAACCCAGTCTCCTCTCCATTTT	2503
QY	2223	AGCCCAATGTTTTCTTGGAGAACCCCTTAATCTTGTCTTCTTGAATTTTAAACCAATT	2282

```

Oy 723 TGGATCCATCTGTCATGCAACCCCTGGTTTACCCCTCAATGAGATGAGAGCTCTTG 782
Db 1004 TGGATCCATCTGTCATGCAACCCCTGGTTTACCCCTCAATGAGATGAGAGCTCTTG 1063
Oy 783 CCAAGATGAGAGAGAGTGTGCCACGAGAACCCCTGGCTGCAAACTGCTCAACACCTA 842
Db 1064 CCAAGATGAGAGAGAGTGTGCCACGAGAACCCCTGGCTGCAAACTGCTCAACACCTA 1123
Oy 843 CGGCTCTTCATCTGCCCTGTGACCCAGATATGAACCTTGAGAGAGAGCTGCTTCAATG 902
Db 1124 CGGCTCTTCATCTGCCCTGTGACCCAGATATGAACCTTGAGAGAGAGCTGCTTCAATG 1183
Oy 903 CAGTATATGAGAGAGTGTGCTGCTGATGCTTCTGCAACATGATGATGATGAGCA 962
Db 1184 CAGTATATGAGAGAGTGTGCTGCTGATGCTTCTGCAACATGATGATGATGAGCA 1243
Oy 963 GCCCGCACATCTTCTGCTCCCTCCAGGCTACATCCTGCTGATGAGCAACCGAAG 1022
Db 1244 GCCCGCACATCTTCTGCTCCCTCCAGGCTACATCCTGCTGATGAGCAACCGAAG 1303
Oy 1023 CTGCCAAGCATCAACGAATGTGACAGAGAACACAGTGTGCAACCTGACAGACAGCTG 1082
Db 1304 CTGCCAAGCATCAACGAATGTGACAGAGAACACAGTGTGCAACCTGACAGACAGCTG 1363
Oy 1083 CTACAATTTACAAGGGGGCTTCAAAATGATCGAACCCCATCGCTGTGAGAGAGCTTATCT 1142
Db 1364 CTACAATTTACAAGGGGGCTTCAAAATGATCGAACCCCATCGCTGTGAGAGAGCTTATCT 1423
Oy 1143 GAGATCACTGATATACCCGCTGATGTGTCTGCTGAGAACCCCTGCTGAGAGAGACAGCC 1202
Db 1424 GAGATCACTGATATACCCGCTGATGTGTCTGCTGAGAACCCCTGCTGAGAGAGACAGCC 1483
Oy 1203 CTTTACCATCTGTACCCGGGACATGAGACGCTGCTGAGAGAGCTCCCTGCTGAGAT 1262
Db 1484 CTTTACCATCTGTGTACCCGGGACATGAGACGCTGCTGAGAGAGCTCCCTGCTGAGAT 1543
Oy 1263 CTTTCCAAATGCAAGCCACGACCCGCTACCTGCGGGCTATTTACATTTTCCAGATCAATC 1322
Db 1544 CTTTCCAAATGCAAGCCACGACCCGCTACCTGCGGGCTATTTACATTTTCCAGATCAATC 1603
Oy 1323 TEGGAATGAGGGCAGAGAAATTTTACATGCGGCAAAAGGGGCCCATCAGTGGCACCTGCT 1382
Db 1604 TEGGAATGAGGGCAGAGAAATTTTACATGCGGCAAAAGGGGCCCATCAGTGGCACCTGCT 1663
Oy 1383 GATGACAGCCCCCATCAAGGGGCCCGGGAATCCAGCTGAGACTTGGAAAATGATCACCTGT 1442
Db 1664 GATGACAGCCCCCATCAAGGGGCCCGGGAATCCAGCTGAGACTTGGAAAATGATCACCTGT 1723
Oy 1443 CAACACTGTCACTCACTTCAAGAGGAGCTCCGTATCCGACTGCGGATATATGTGTGCGA 1502
Db 1724 CAACACTGTCACTCACTTCAAGAGGAGCTCCGTATCCGACTGCGGATATATGTGTGCGA 1783
Oy 1503 GTACCACTTCTGAGCCTCGGGCTGAGAGCTCCGACGCTGCTCTCATTTGACCAAGGGA 1562
Db 1784 GTACCACTTCTGAGCCTCGGGCTGAGAGCTCCGACGCTGCTCTCATTTGACCAAGGGA 1843
Oy 1563 CAGGAGAGAGAGAGAAATTAACAGAGAGAAATGAGAGGACACAGAGCTTAGGCAATTCCTG 1622
Db 1844 CAGGAGAGAGAGAGAAATTAACAGAGAGAAATGAGAGGACACAGAGCTTAGGCAATTCCTG 1903
Oy 1623 CTGAAGCTTTCCCGGAGAGATGAGCCCGACTTCTGACTCTCACTGATATTTGAGCA 1682
Db 1904 CTGAAGCTTTCCCGGAGAGATGAGCCCGACTTCTGACTCTCACTGATATTTGAGCA 1963
Oy 1683 CCGTGCACCCCTGACAGAGCTTCCACCCAGTTCTATGATATCAAGGCTTCAAGAAAGTAT 1742
Db 1964 CCGTGCACCCCTGACAGAGCTTCCACCCAGTTCTATGATATCAAGGCTTCAAGAAAGTAT 2023
Oy 1743 ATCATTGCTCCCTGATTAAGATTTGTGTGATTAATTTCAAGGCTTCAAGTATTTTCCA 1802
Db 2024 ATCATTGCTCCCTGATTAAGATTTGTGTGATTAATTTCAAGGCTTCAAGTATTTTCCA 2083

```

```

Oy 1803 CTAATTTCAAGAAATAGATAGATTGCGGGGGTCTGAGCTATGTTCAAGAGACTG 1862
Db 2084 CTAATTTCAAGAAATAGATAGATTGCGGGGGTCTGAGCTATGTTCAAGAGACTG 2143
Oy 1863 AACAGCTTGTCTGACTTCTTCAACCTTCCACTCTCTCTCACTGTTACTGCTT 1922
Db 2144 AACAGCTTGTCTGACTTCTTCAACCTTCCACTCTCTCTCACTGTTACTGCTT 2203
Oy 1923 CAAAGACCCGGGAGCTGCGGGGAAACCTGGGAGTAGTACTGTTGCTTTGCGTAC 1982
Db 2204 CAAAGACCCGGGAGCTGCGGGGAAACCTGGGAGTAGTACTGTTGCTTTGCGTAC 2263
Oy 1983 GAGAGAGCTATGTAACAAACACAGAGATGCAAGAGTTTGAAGAATGTTTCA 2042
Db 2264 GAGAGAGCTATGTAACAAACACAGAGATGCAAGAGTTTGAAGAATGTTTCA 2323
Oy 2043 AACCATGCTGTGATTTTCAACCATTAAGAGATTGAGTGTCTTAAATTTGATTAAC 2102
Db 2324 AACCATGCTGTGATTTTCAACCATTAAGAGATTGAGTGTCTTAAATTTGATTAAC 2383
Oy 2103 GGTTAATCTGTCTGTCTGATTTGATTTTAAATAATGCTAGATTCCTTC 2162
Db 2384 GGTTAATCTGTCTGTCTGATTTGATTTTAAATAATGCTAGATTCCTTC 2443
Oy 2163 GAAAGCCCTTCAACACATGATGTTCTGTCTTCCAAACCACTCTCTCCATTT 2222
Db 2444 GAAAGCCCTTCAACACATGATGTTCTGTCTTCCAAACCACTCTCTCCATTT 2503
Oy 2223 AGCCAGTGTCTTCTTGAAGACCCCTTAATCTGCTTCTTTAGATTTTACCAAT 2282
Db 2504 AGCCAGTGTCTTCTTGAAGACCCCTTAATCTGCTTCTTTAGATTTTACCAAT 2563
Oy 2283 GATGGAATGACAGAGTCTCCAAACGATTAATTTTGAAGGA 2328
Db 2564 GATGGAATGACAGAGTCTCCAAACGATTAATTTTGAAGGA 2609

```

RESULT 9

US-10-175-738-407

Sequence 407, Application US/10175738
 Publication No. US2003022294A1
 GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
 APPLICANT: Chen, Jian
 APPLICANT: Desnoyers, Luc
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Pan, James
 APPLICANT: Smith, Victoria
 APPLICANT: Watanabe, Colin K.
 APPLICANT: Wood, William I.
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 FILE OF INVENTION: ACIDS ENCODING THE SAME
 FILE REFERENCE: P3430R1C45
 CURRENT APPLICATION NUMBER: US/10/175,738
 CURRENT FILING DATE: 2002-06-19
 Prior application removed - See file wrapper or Palm
 NUMBER OF SEQ ID NOS: 612
 SEQ ID NO 407
 TYPE: DNA
 LENGTH: 2609
 ORGANISM: Homo Sapien
 US-10-175-738-407

Query Match 99.8%; Score 2322.8; DB 9; Length 2609;
 Best Local Similarity 99.9%; Pred. No. 0;

Matches 2324; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Oy 3 CCGGCGCTCTCCCGCTGCTCTCCACGAGCTGCTGCGCCCTCTGGAATTAACACCC 62
Db 284 CCGGCGCTCTCCCGCTGCTCTCCACGAGCTGCTGCGCCCTCTGGAATTAACACCC 343

```


|||||
Db 764 CACGATCTCCAGGCGCTTATATGCGCGCTTGGATACGATGGATGAAGCAACCATG 823
QY 543 TGTGGATGTGGACGAGTGTGCAACAGATTTCCACGATGCAACCCACCATGATCTGCAT 602
Db 824 TGTGGATGTGGACGAGTGTGCAACAGATTTCCACGATGCAACCCACCATGATCTGCAT 883
QY 603 CAATCTGTAAGGCGGGGTACACCTGCTCTGCAACGAGATTTGGTTGTGGAAAGGCA 662
Db 884 CAATCTGTAAGGCGGGGTACACCTGCTCTGCAACGAGATTTGGTTGTGGAAAGGCA 943
QY 663 GTGCTTAGACATGTGTAATGTGCGATGTGTTACTGCGACGAGCTCTGTGCGAATGTTCC 722
Db 944 GTGCTTAGACATGTGTAATGTGCGATGTGTTACTGCGACGAGCTCTGTGCGAATGTTCC 1003
QY 723 TGGATCTTATTTTGTACATGCAACCCCTGGTTTACCCTCAATGAGATGGAAGTCTTG 782
Db 1004 TGGATCTTATTTTGTACATGCAACCCCTGGTTTACCCTCAATGAGATGGAAGTCTTG 1063
QY 783 CCAAGATGTAAGGAGTGTGCGACGAGAACCCCTGGTGCACAACTGGCGTCAACACCTA 842
Db 1064 CCAAGATGTAAGGAGTGTGCGACGAGAACCCCTGGTGCACAACTGGCGTCAACACCTA 1123
QY 843 CGGCTCTTTCATCTGCGCTGTGACCCAGATATGAATTGAGGAAGATGGCGTTCAATG 902
Db 1124 CGGCTCTTTCATCTGCGCTGTGACCCAGATATGAATTGAGGAAGATGGCGTTCAATG 1183
QY 903 CAGTATATGACGAGTGTGCAAGCTTCTGTAGTTCCTTGGCAACATGATGTGTAACCA 962
Db 1184 CAGTATATGACGAGTGTGCAAGCTTCTGTAGTTCCTTGGCAACATGATGTGTAACCA 1243
QY 963 GCCCGGACATCTTGTGCTGCTGCGCTCCAGGCTACATCCTGCTGATGACCAACCGAG 1022
Db 1244 GCCCGGACATCTTGTGCTGCTGCGCTCCAGGCTACATCCTGCTGATGACCAACCGAG 1303
QY 1023 CTGCCAAGACATCAACGAATGTGAGACAGAAACACAGCTGCAACCTGACAGACGTG 1082
Db 1304 CTGCCAAGACATCAACGAATGTGAGACAGAAACACAGCTGCAACCTGACAGACGTG 1363
QY 1083 CTACATATTTACAAGGGGGCTTCAAAATGATCAGACCCCATCCGCTGTAGAGAGCTTATCT 1142
Db 1364 CTACATATTTACAAGGGGGCTTCAAAATGATCAGACCCCATCCGCTGTAGAGAGCTTATCT 1423
QY 1143 GAGGATCAGTGAATTAACCGCTGTATGTCTGCTGAGAACCCCTGGCTGACAGACAGCC 1202
Db 1424 GAGGATCAGTGAATTAACCGCTGTATGTCTGCTGAGAACCCCTGGCTGACAGACAGCC 1483
QY 1203 CTTTACCATCTTGTACCGGAGACATGACGAGTGTGTCAGAGCGTCCGTTCCCGTGACAT 1262
Db 1484 CTTTACCATCTTGTACCGGAGACATGACGAGTGTGTCAGAGCGTCCGTTCCCGTGACAT 1543
QY 1263 CTTCCAAATGCAAGCCACGACCCGCTACCTGGGGCTTATTAATTTTCCAGATCAATC 1322
Db 1544 CTTCCAAATGCAAGCCACGACCCGCTACCTGGGGCTTATTAATTTTCCAGATCAATC 1603
QY 1323 TGGGATGAGGGGAGAGAAATTTTACATGCGGCAAAAGGGGGCCCATCAGTGCACCCCTG 1382
Db 1604 TGGGATGAGGGGAGAGAAATTTTACATGCGGCAAAAGGGGGCCCATCAGTGCACCCCTG 1663
QY 1383 GATGACACGCCCCCATCAAGGGGGCCCGGGAATCCAGCTGGACTTGGAAATGATCACTGT 1442
Db 1664 GATGACACGCCCCCATCAAGGGGGCCCGGGAATCCAGCTGGACTTGGAAATGATCACTGT 1723
QY 1443 CAACACTGTCAATCACTTCAAGAGGAGCTCCGATCCGAGCTGGGATATATGTGTGCA 1502
Db 1724 CAACACTGTCAATCACTTCAAGAGGAGCTCCGATCCGAGCTGGGATATATGTGTGCA 1783
QY 1503 GTACCATTTCTGAGCTCGGGCTGAGAGCTCCGAGCGTGCCTTCATTTGGACCAAGGGA 1562
Db 1784 GTACCATTTCTGAGCTCGGGCTGAGAGCTCCGAGCGTGCCTTCATTTGGACCAAGGGA 1843
QY 1563 CAGGAGAGAGAGAAATACAGAGAGAAATGAGAGCAGACAGAGCTTGGCATTTCTGT 1622
|||||

Db 1844 CAGGAGAGAGAGAAATACAGAGAGAAATGAGAGCAGACAGAGCTTGGCATTTCTG 1903
QY 1623 CTGACGCTTTCOCAGAGAGTACGCCCGGCTTCCTGACTCTCAGTACTATTTGAGA 1682
Db 1904 CTGACGCTTTCOCAGAGAGTACGCCCGGCTTCCTGACTCTCAGTACTATTTGAGA 1963
QY 1683 CCTGTACCCCTGAGAGACTTGCACCCCGGAGTTCCTATGATACAGTATCAAAAAGTAT 1742
Db 1964 CCTGTACCCCTGAGAGACTTGCACCCCGGAGTTCCTATGATACAGTATCAAAAAGTAT 2023
QY 1743 ATCATTTGCTCCCTGATAGAAAGTGTGTGTAATTTTCAAGGCGCTTCACTTATTTCCA 1802
Db 2024 ATCATTTGCTCCCTGATAGAAAGTGTGTGTAATTTTCAAGGCGCTTCACTTATTTCCA 2083
QY 1803 CTATTTTCAAGAAATAGATTTAGTTTGGGGGGCTCGAGTCTATGTTCAAGAGCTGTG 1862
Db 2084 CTATTTTCAAGAAATAGATTTAGTTTGGGGGGCTCGAGTCTATGTTCAAGAGCTGTG 2143
QY 1863 AACAGCTGCTGCTACCTTCTTCAACCTTCCACTCTCTCTCACTGTGTACTGCTTGTG 1922
Db 2144 AACAGCTGCTGCTACCTTCTTCAACCTTCCACTCTCTCTCACTGTGTACTGCTTGTG 2203
QY 1923 CAAAGACCCGGAGAGTGTGGGGGGAACCCCTGGAGTACCTGATTTGCTTTGGCGTACA 1982
Db 2204 CAAAGACCCGGAGAGTGTGGGGGGAACCCCTGGAGTACCTGATTTGCTTTGGCGTACA 2263
QY 1983 GAGAAGCTATGTAAACAACAACAGAGAGATCGAAGGGTTTTAGAGATGTGTTCAA 2042
Db 2264 GAGAAGCTATGTAAACAACAACAGAGAGATCGAAGGGTTTTAGAGATGTGTTCAA 2323
QY 2043 AACCATGCTGTGATTTTCAACCATTAAGAAAGTTTCAAGTTGCTTAAATTTGTATAC 2102
Db 2324 AACCATGCTGTGATTTTCAACCATTAAGAAAGTTTCAAGTTGCTTAAATTTGTATAC 2383
QY 2103 GCTTATATCTGTCTGTATTTTGTAGTATTTTAAATAATAGTCTTAATTCCTTC 2162
Db 2384 GCTTATATCTGTCTGTATTTTGTAGTATTTTAAATAATAGTCTTAATTCCTTC 2443
QY 2163 GAAAGGCTTCAGACACATGCTATGTTCTGCTTCCCAACCCAGTCTCTCTCATTTT 2222
Db 2444 GAAAGGCTTCAGACACATGCTATGTTCTGCTTCCCAACCCAGTCTCTCTCATTTT 2503
QY 2223 AGCCAGTGTCTTCTTGAAGACCCCTTAATCTTCTTCTTGAATTTTACCATTT 2282
Db 2504 AGCCAGTGTCTTCTTGAAGACCCCTTAATCTTCTTCTTGAATTTTACCATTT 2563
QY 2283 GCATTGGAATGACAGAGTCTTCAACAGTGAATTAATTTTGAAGAGA 2328
Db 2564 GCATTGGAATGACAGAGTCTTCAACAGTGAATTAATTTTGAAGAGA 2609
|||||

RESULT 6
US-10-176-758-407
; Sequence 407, Application US/10176758
; Publication No. US2003008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C104
; CURRENT APPLICATION NUMBER: US/10/176,758
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm

```
Db 1424 GAGGATGAGTAAACCGCTGATGTGCTGCTGAGAACCTGGCTGAGAGACGACC 1483
QY 1203 CTTTACCATCTTGACCGGGGACATGAGAGTGTGTCAGAGACCTCCGTTCCCGCTACAT 1262
Db 1484 CTTTACCATCTTGACCGGGGACATGAGAGTGTGTCAGAGACCTCCGTTCCCGCTACAT 1543
QY 1263 CTTTCAAAATGCAAGCCAGACCCGCTACCTGGGGCTATTTCATTTTCCAGTCAAAATC 1322
Db 1544 CTTTCAAAATGCAAGCCAGACCCGCTACCTGGGGCTATTTCATTTTCCAGTCAAAATC 1603
QY 1323 TGGGAATGAGGGGCGAGAAATTTTACATGCGGCAAAAGGGCCCAATGTCACCCCTGT 1382
Db 1604 TGGGAATGAGGGGCGAGAAATTTTACATGCGGCAAAAGGGCCCAATGTCACCCCTGT 1663
QY 1383 GATGACACGGCCCATCAAAAGGGCCCGGGAATTCAGCTGAGCTTGGAATGATCAGCT 1442
Db 1664 GATGACACGGCCCATCAAAAGGGCCCGGGAATTCAGCTGAGCTTGGAATGATCAGCT 1723
QY 1443 CAACACTGTCATCAACTTTCAGAGGAGCTCCGTGATCCGACTGCGGATATATGTGCGCA 1502
Db 1724 CAACACTGTCATCAACTTTCAGAGGAGCTCCGTGATCCGACTGCGGATATATGTGCGCA 1783
QY 1503 GTTACCATTTTGAGCTTGGGCTGGAGCTTCGACGCTGCTCTGATTTGGCAACCAAGGA 1562
Db 1784 GTTACCATTTTGAGCTTGGGCTGGAGCTTCGACGCTGCTCTGATTTGGCAACCAAGGA 1843
QY 1563 CAGGAGAGAGAGAAATTAACAGAGAAATGAGAGGACGACAGAGAGTTGGATTTCCGG 1622
Db 1844 CAGGAGAGAGAGAAATTAACAGAGAAATGAGAGGACGACAGAGAGTTGGATTTCCGG 1903
QY 1623 CTGAACTTTTCCCGGAGAGTACAGCCCGCTGCTGACTCTCACCTGACTATTGCGAGA 1682
Db 1904 CTGAACTTTTCCCGGAGAGTACAGCCCGCTGCTGACTCTCACCTGACTATTGCGAGA 1963
QY 1683 CCGTGTACCTGCGAGGACTTCCACCCAGTCTCTATGATATACGTTTCAAAAAGTAT 1742
Db 1964 CCGTGTACCTGCGAGGACTTCCACCCAGTCTCTATGATATACGTTTCAAAAAGTAT 2023
QY 1743 ATGATGCTCCCGGAGAGAAATTTGTTGTAATTTTCAAGGCTTCGTTGTTATTTCA 1802
Db 2024 ATGATGCTCCCGGAGAGAAATTTGTTGTAATTTTCAAGGCTTCGTTGTTATTTCA 2083
QY 1803 CTATTTTCAAGAAATAGATAGTTAGTTGCGGGGCTGAGTCTATGTTCAAAAGACTGT 1862
Db 2084 CTATTTTCAAGAAATAGATAGTTAGTTGCGGGGCTGAGTCTATGTTCAAAAGACTGT 2143
QY 1863 AAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1922
Db 2144 AAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2203
QY 1923 CAAGAAGCCCGGAGAGTGGGGGAAACCTGGGAGTAGTTGCTTTTCCGTTGCGACA 1982
Db 2204 CAAGAAGCCCGGAGAGTGGGGGAAACCTGGGAGTAGTTGCTTTTCCGTTGCGACA 2263
QY 1983 GAGAAGGCTATGTAACAAACACAGACAGATCGAAGGGTTTAAAGAAATGTTTCA 2042
Db 2264 GAGAAGGCTATGTAACAAACACAGACAGATCGAAGGGTTTAAAGAAATGTTTCA 2323
QY 2043 AACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2102
Db 2324 AACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2383
QY 2103 GCTTTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2162
Db 2384 GCTTTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2443
QY 2163 GAAAGGCTTACAGACATGCTATGTTGCTTCCAAACCCAGTCTCTCTCCATTTT 2222
Db 2444 GAAAGGCTTACAGACATGCTATGTTGCTTCCAAACCCAGTCTCTCTCCATTTT 2503
QY 2223 AGCCCATGTTTCTTGGAGACCCCTTAATCTGCTTCTTGAATTTTACCAANT 2282
Db 2504 AGCCCATGTTTCTTGGAGACCCCTTAATCTGCTTCTTGAATTTTACCAANT 2563
QY 2283 GGATTGGAATGACAGGCTTCCAAACCTGATTAATTTGAAGAGA 2328
Db 2564 GGATTGGAATGACAGGCTTCCAAACCTGATTAATTTGAAGAGA 2609

RESULT 5
US-10-174-590-407
: Sequence 407, Application US/10174590
: Publication No. US2003008352A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Chen, Jian
: APPLICANT: Desnoyers, Luc
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Pan, James
: APPLICANT: Smith, Victoria
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P3430R1CA2
: CURRENT APPLICATION NUMBER: US/10/174,590
: PRIORITY FILING DATE: 2002-06-18
: Prior application removed - See File Wrapper or Palm
: NUMBER OF SEQ ID NOS: 612
: SEQ ID NO 407
: LENGTH: 2609
: TYPE: DNA
: ORGANISM: Homo Sapien
US-10-174-590-407

Query Match 99.8%; Score 2322.8; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2324; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CCGCGGCTGCTCCCGTGTCTCTCTGACGACTGCTGCGCCCTCTGGAATAAAACACC 62
Db 284 CCGCGGCTGCTCCCGTGTCTCTCTGACGACTGCTGCGCCCTCTGGAATAAAACACC 343
QY 63 GCGAGCCCGGAGGCGGAGAGGCGAGCGAGTCCCGGAGTCTCCGCGGGGTCCCGCG 122
Db 344 GCGAGCCCGGAGGCGGAGAGGCGAGCGAGTCCCGGAGTCTCCGCGGGGTCCCGCG 403
QY 123 CGAGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 182
Db 404 CGAGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 463
QY 183 AAGATCTCAGTGTATTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 242
Db 464 AAGATCTCAGTGTATTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 523
QY 243 AAGATCTCAGTGTATTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302
Db 524 AAGATCTCAGTGTATTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 583
QY 303 ATGCGGAACATCCCGGAGGCTGCGGAGAGACATGATGTGTTAAACCAAAATGCGCG 362
Db 584 ATGCGGAACATCCCGGAGGCTGCGGAGAGACATGATGTGTTAAACCAAAATGCGCG 643
QY 363 GTATTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422
Db 644 GTATTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 703
QY 423 GACCCCTTACTCAGATCCGTAACCAAGAGCTGCCCCACACATCTCAGTCCAAATATTC 482
Db 704 GACCCCTTACTCAGATCCGTAACCAAGAGCTGCCCCACACATCTCAGTCCAAATATTC 763
QY 483 CAGCATCTCCAGGCTCTTATATGCGCGCTTGGATACCAAGATGATGAAAGCAACCAATG 542
```


PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380138
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380139
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/403296
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/403297
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/423741
PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: 09/423844
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 09/522342
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: 09/548815
PRIOR FILING DATE: 2000-04-13
PRIOR APPLICATION NUMBER: 09/664610
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: 09/665350
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: 09/709238
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 09/767609
PRIOR FILING DATE: 2001-01-22
PRIOR APPLICATION NUMBER: 09/802706
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 09/808689
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 09/866028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 09/870574
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: 09/872035
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 09/886342
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: PCT/US98/14552
PRIOR FILING DATE: 1998-07-14
PRIOR APPLICATION NUMBER: PCT/US98/18824
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: PCT/US98/19093
PRIOR FILING DATE: 1998-09-14
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: PCT/US98/19437
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: PCT/US98/24855
PRIOR FILING DATE: 1998-11-20
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: 1998-12-01
PRIOR APPLICATION NUMBER: PCT/US98/25190
PRIOR FILING DATE: 1998-11-25
PRIOR APPLICATION NUMBER: PCT/US99/05028
PRIOR FILING DATE: 1999-03-08
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/20111
PRIOR FILING DATE: 1999-09-01
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547

Query Match 99.8%; Score 2322.8; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2324; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CCGGCGCTCTCCCGTGTCTCTCCACGACGCTGGCTGGCCCTCTGGAATAAACACCC 62
DB 284 CCGGCGCTCTCCCGTGTCTCTCTCCACGACGCTGGCCCTCTGGAATAAACACCC 343

QY 63 GCGAGCCCGGAGGCGCCAGAGAGGCGGAGCTGCGCCGAGCTCTCCGGGGGTCGCCCGC 122
DB 344 GCGAGCCCGGAGGCGCCAGAGAGGCGGAGCTGCGCCGAGCTCTCCGGGGGTCGCCCGC 403
QY 123 CGAGCTTTCTTCTCGCCTTTCGATCTCTCTCTCGCGCTTGTGACATGCGAGAAATAA 182
DB 404 CGAGCTTTCTTCTCGCCTTTCGATCTCTCTCTCGCGCTTGTGACATGCGAGAAATAA 463
QY 183 AAGGATACCTGTTTACATCTGCGCTCTCTCTCTCAAGCCCTGGGAATGACAGGC 242
DB 464 AAGGATACCTGTTTACATCTGCGCTCTCTCTCTCAAGCCCTGGGAATGACAGGC 523
QY 243 ACAGTGCAGAAATGCTTGTGACCTGATGCGCAGTGCAGCAGTGTATATATATGA 302
DB 524 ACAGTGCAGAAATGCTTGTGACCTGATGCGCAGTGCAGCAGTGTATATATATGA 583
QY 303 ATGCCGAACATCCCGGAGGCTGCGGAGAGACATGATGTGTATTAACCAATGGCGG 362
DB 584 ATGCCGAACATCCCGGAGGCTGCGGAGAGACATGATGTGTATTAACCAATGGCGG 643
QY 363 GTATTATGCAATTCGCCGACAAACCTGTGTATGAGGCGCTACTGAAACCCCTACTC 422
DB 644 GTATTATGCAATTCGCCGACAAACCTGTGTATGAGGCGCTACTGAAACCCCTACTC 703
QY 423 GACCCCTACTAGTCCGTACCCAGACAGCTGCCACACACTCTCAGCTCCAAATATCC 482
DB 704 GACCCCTACTAGTCCGTACCCAGACAGCTGCCACACACTCTCAGCTCCAAATATCC 763
QY 483 CACGATCTCAGGCTCTTATATGCGCTTGTGATACCAATGATGATTAAGCAACATG 542
DB 764 CACGATCTCAGGCTCTTATATGCGCTTGTGATACCAATGATGATTAAGCAACATG 823
QY 543 TGTGATGTGAGAGAGTGTGCAACAGATTCACACAGTGCACCCACCAAGATCTGCAT 602
DB 824 TGTGATGTGAGAGAGTGTGCAACAGATTCACACAGTGCACCCACCAAGATCTGCAT 883
QY 603 CAATAGTGAAGCGGCTTACCTCTCTCTGACCGAGATGATGCTTCTGGAAGGCA 662
DB 884 CAATAGTGAAGCGGCTTACCTCTCTCTGACCGAGATGATGCTTCTGGAAGGCA 943
QY 663 GTGCTTAGCAATGATGATGATGCTATGCTTGTGCTGACAGCTGTGGAATGTTC 722
DB 944 GTGCTTAGCAATGATGATGATGCTATGCTTGTGCTGACAGCTGTGGAATGTTC 1003
QY 723 TEGATCTTATCTTGTGATGCAACCTGCTTTTACCCTCAATGAGATGGAAGTCTTG 782
DB 1004 TEGATCTTATCTTGTGATGCAACCTGCTTTTACCCTCAATGAGATGGAAGTCTTG 1063
QY 783 CCAAGATGTGAACGAGTGTGCAACCGAAGACCTGCTGTAACCTGCTCAACACTA 842
DB 1064 CCAAGATGTGAACGAGTGTGCAACCGAAGACCTGCTGTAACCTGCTCAACACTA 1123
QY 843 CGGCTCTTCAATGCGGCTGTGACCGAGATGATGAGAAATGAGCTGTCTATG 902
DB 1124 CGGCTCTTCAATGCGGCTGTGACCGAGATGATGAGAAATGAGCTGTCTATG 1183
QY 903 CAGTGTATGAGAGAGTGTGCAAGTGTGCTGCTGCTGCAACATGAGTGTGAACCA 962
DB 1184 CAGTGTATGAGAGAGTGTGCAAGTGTGCTGCTGCTGCAACATGAGTGTGAACCA 1243
QY 963 GCGCGGACATCTTCTGCTCTGCGCTCCAGAGTATCTGCTGTGATGACACCGAAG 1022
DB 1244 GCGCGGACATCTTCTGCTCTGCGCTCCAGAGTATCTGCTGTGATGACACCGAAG 1303
QY 1023 CTGCGCAAGATGAGAAATGTGAGCAGAGAACACAGTGCAGCTGCAGACAGCTG 1082
DB 1304 CTGCGCAAGATGAGAAATGTGAGCAGAGAACACAGTGCAGCTGCAGACAGCTG 1363
QY 1083 CTACAAATTTACAGAGGCTTCAAAATGATGACCCCTGCTGTGAGAGCCCTATCT 1142
DB 1364 CTACAAATTTACAGAGGCTTCAAAATGATGACCCCTGCTGTGAGAGCCCTATCT 1423
QY 1143 GAGGATCAATGATTAACCGCTGTATGTGTCTCTGTGAGAACCTGGCTGCAGAGACAGCC 1202

Db 2436 AGCCAGGTTTCTTTAGGACCCCTTAATCTTCTTTAGATTTTACCAAT 2495
Oy 2283 GGATTGGAATGACAGAGTCTCCAACTGATTAATATTTGAGAGA 2328
Db 2496 GGATTGGAATGACAGAGTCTCCAACTGATTAATATTTGAGAGA 2541

RESULT 4

US-10-066-500-14
Sequence 14, Application US/10066500
Patent No. US20020177165A1
GENERAL INFORMATION:
APPLICANT: Avi J. Ashkenazi
APPLICANT: Kevin P. Baker
APPLICANT: David A. Botstein
APPLICANT: Luc Desnoyers
APPLICANT: Dan L. Eaton
APPLICANT: Napoleone Ferrara
APPLICANT: Sherman Fong
APPLICANT: Wei-Qiang Gao
APPLICANT: Hanspeter Gerber
APPLICANT: Mary E. Gertlisen
APPLICANT: Audrey Goddard
APPLICANT: Paul J. Godowski
APPLICANT: Austin L. Gurney
APPLICANT: Ivar J. Kljavin
APPLICANT: Jennie P. Mather
APPLICANT: Mary A. Napier
APPLICANT: James Pan
APPLICANT: Nicholas F. Paoni
APPLICANT: Margaret Ann Roy
APPLICANT: Timothy A. Stewart
APPLICANT: Daniel Tumas
APPLICANT: Colin K. Watanabe
APPLICANT: P. Mickey Williams
APPLICANT: William T. Wood
APPLICANT: Zemin Zang
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P313081C7
CURRENT APPLICATION NUMBER: US/10/066,500
CURRENT FILING DATE: 2002-02-01
PRIOR APPLICATION NUMBER: 10/002,796
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/062285
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062816
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063082
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/063329
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063733
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063664
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066840
PRIOR FILING DATE: 1997-11-25
PRIOR APPLICATION NUMBER: 60/069694
PRIOR FILING DATE: 1997-12-16
PRIOR APPLICATION NUMBER: 60/074086
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/074092
PRIOR FILING DATE: 1998-02-09

PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/095998
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/097000
PRIOR FILING DATE: 1998-08-18
PRIOR APPLICATION NUMBER: 60/099601
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099811
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100858
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101922
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/106032
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: 60/109304
PRIOR FILING DATE: 1998-11-20
PRIOR APPLICATION NUMBER: 60/125778
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/139695
PRIOR FILING DATE: 1999-06-15
PRIOR APPLICATION NUMBER: 60/145070
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/149396
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/169495
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 08/918874
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 08/933821
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 08/960507
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 09/114844
PRIOR FILING DATE: 1998-07-14
PRIOR APPLICATION NUMBER: 09/136801
PRIOR FILING DATE: 1998-08-19
PRIOR APPLICATION NUMBER: 09/136804
PRIOR FILING DATE: 1998-08-19
PRIOR APPLICATION NUMBER: 09/136828
PRIOR FILING DATE: 1998-08-19
PRIOR APPLICATION NUMBER: 09/158342
PRIOR FILING DATE: 1998-09-21
PRIOR APPLICATION NUMBER: 09/180997
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 09/202088
PRIOR FILING DATE: 1998-12-08
PRIOR APPLICATION NUMBER: 09/254311
PRIOR FILING DATE: 1999-03-03
PRIOR APPLICATION NUMBER: 09/254460
PRIOR FILING DATE: 1999-03-09
PRIOR APPLICATION NUMBER: 09/254465
PRIOR FILING DATE: 1999-03-05
PRIOR APPLICATION NUMBER: 09/284663
PRIOR FILING DATE: 1999-04-15
PRIOR APPLICATION NUMBER: 09/332928
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/332929
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/333075
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/333077
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/380137

QY	63	GGGAGCCCGAGGGCCGAGAGGCGAGCGGCCGAGCTCTCCGGGGGTC	122
Db	276	GGAGGCCCGAGGGCCGAGAGGCGAGCGGCCGAGCTCTCCGGGGGTC	335
QY	123	CGAGCTTCTCTGCGCTTCGATCTCTCTCGCGCGCTCTTGAGATCCAGATAA	182
Db	336	CGAGCTTCTCTGCGCTTCGATCTCTCTCGCGCGCTCTTGAGATCCAGATAA	395
QY	183	AAGATATCTACCTTTACATCTGCGCTCTCTCTCTCCAGCCCTGGGAATGACAGGC	242
Db	396	AAGATATCTACCTTTACATCTGCGCTCTCTCTCTCCAGCCCTGGGAATGACAGGC	455
QY	243	ACAGTGCAGCAATGGCTTTGACCTGAGTCCGAGTCCAGAGACAGTGTATGATGATGA	302
Db	456	ACAGTGCAGCAATGGCTTTGAGCTGAGTCCGAGTCCAGAGACAGTGTATGATGATGA	515
QY	303	ATGCGCAACATCCCGAGGCTCCGAGAGACATGATGTGTAAACCAAAATGGCGG	362
Db	516	ATGCGCAACATCCCGAGGCTCCGAGAGACATGATGTGTAAACCAAAATGGCGG	575
QY	363	GTATTTATGATTCGCGGAGAAACCCGTGTATGAGAGGCCCTACTGAAACCTACTC	422
Db	576	GTATTTATGATTCGCGGAGAAACCCGTGTATGAGAGGCCCTACTGAAACCTACTC	635
QY	423	GACCCCTACTCAGAGTCCGTAACCCAGAGCTGCCACACACTCTCAGCTCCAACTATCC	482
Db	636	GACCCCTACTCAGAGTCCGTAACCCAGAGCTGCCACACACTCTCAGCTCCAACTATCC	695
QY	483	CACGATCTCCAGGCTCTTATATGCGCTTTGATGATCAGATGATGAAAGCAACCAATG	542
Db	696	CACGATCTCCAGGCTCTTATATGCGCTTTGATGATCAGATGATGAAAGCAACCAATG	755
QY	543	TGTGATGTGAGAGTGTGCAACAGATTCACAGTGCAGAACCCACCCAGATGCTGAT	602
Db	756	TGTGATGTGAGAGTGTGCAACAGATTCACAGTGCAGAACCCACCCAGATGCTGAT	815
QY	603	CAATACTGAAGGGGGTACACCTGCTCTGACACGAGATATGGCTTCTGGAAGGCCA	662
Db	816	CAATACTGAAGGGGGTACACCTGCTCTGACACGAGATATGGCTTCTGGAAGGCCA	875
QY	663	GTCCTTAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	722
Db	876	GTCCTTAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	935
QY	723	TGATCTTATCTTGTACATGCAACCCGTTTACCTCAATGATGATGATGATGATGATG	782
Db	936	TGATCTTATCTTGTACATGCAACCCGTTTACCTCAATGATGATGATGATGATGATG	995
QY	783	CCAAGATGTGAAGAGTGTGCGACCGAGAACCCCTGCGTGCACAACTGTCACACCTA	842
Db	996	CCAAGATGTGAAGAGTGTGCGACCGAGAACCCCTGCGTGCACAACTGTCACACCTA	1055
QY	843	CGGCTCTTATCTGCGCGCTGTGACCGAGATATGAACTTGAGGAAGATGGCGTTTATG	902
Db	1056	CGGCTCTTATCTGCGCGCTGTGACCGAGATATGAACTTGAGGAAGATGGCGTTTATG	1115
QY	903	CAGTATATGAGAGTGCAGCTCTCTGAGTTCCTCTGCAACATGATGATGATGATGATG	962
Db	1116	CAGTATATGAGAGTGCAGCTCTCTGAGTTCCTCTGCAACATGATGATGATGATGATG	1175
QY	963	GCCCGGCAATCTTCTGCTCTGCGCTCCAGGCTACATCTGCTGATGACAAACCGAAG	1022
Db	1176	GCCCGGCAATCTTCTGCTCTGCGCTCCAGGCTACATCTGCTGATGACAAACCGAAG	1235
QY	1023	CTGCCAAGACATACAGAAATGTGAGCAGAGAACCAACAGCTGCAACCTGCAAGAGCTG	1082
Db	1236	CTGCCAAGACATACAGAAATGTGAGCAGAGAACCAACAGCTGCAACCTGCAAGAGCTG	1295
QY	1083	CTACATTTTCAAGAGGGGCTTCAAAATGATGATGATGATGATGATGATGATGATGATG	1142
Db	1296	CTACATTTTCAAGAGGGGCTTCAAAATGATGATGATGATGATGATGATGATGATGATG	1355
QY	1143	GAGATACATGATACACCGCTGATGTGCTGCTGAGAACCTGCGTCAAGACCAAGCC	1202
Db	1356	GAGATACATGATACACCGCTGATGTGCTGCTGAGAACCTGCGTCAAGACCAAGCC	1415
QY	1203	CTTTACCATCTTGTACCGGGGACATGAGAGTGTGACAGAGCTCCGTTCCGCTGATAC	1262
Db	1416	CTTTACCATCTTGTACCGGGGACATGAGAGTGTGACAGAGCTCCGTTCCGCTGATAC	1475
QY	1263	CTTCCAAATGCAAGCCGACAGCCGCTACCTGCGGGCTATTACATTTTCCAGATCAATC	1322
Db	1476	CTTCCAAATGCAAGCCGACAGCCGCTACCTGCGGGCTATTACATTTTCCAGATCAATC	1535
QY	1323	TGGGAATGAGGGGAGAAATTTTACATGCGGCAAAAGGGGCCCATGAGTCCAGCTGAT	1382
Db	1536	TGGGAATGAGGGGAGAAATTTTACATGCGGCAAAAGGGGCCCATGAGTCCAGCTGAT	1595
QY	1383	GATGACAGCCCATCAAGAGGCCCCGGGAAATCCAGCTGAGCTTGAATGATCAGTGT	1442
Db	1596	GATGACAGCCCATCAAGAGGCCCCGGGAAATCCAGCTGAGCTTGAATGATCAGTGT	1655
QY	1443	CAACACTGTATCAACTTCAGAGAGCAGCTCCGATCCGATGCGGGATATATGTGCGCA	1502
Db	1656	CAACACTGTATCAACTTCAGAGAGCAGCTCCGATCCGATGCGGGATATATGTGCGCA	1715
QY	1503	GTACCCATCTGAGGCTCGGGGCTGAGGCTCCGAGCGCTGCTCATGATGAGCAAGGGA	1562
Db	1716	GTACCCATCTGAGGCTCGGGGCTGAGGCTCCGAGCGCTGCTCATGATGAGCAAGGGA	1775
QY	1563	CAGGAGAGAGAGAAATTAACAGAGAAATGAGAGCAGACAGACGTTAGGCATTTCTG	1622
Db	1776	CAGGAGAGAGAGAAATTAACAGAGAAATGAGAGCAGACAGACGTTAGGCATTTCTG	1835
QY	1623	CTGAGCTTCCCGAAGAGTGCAGCCCGACTTCTGCTGCTGCTGCTGCTGCTGCTGCTG	1682
Db	1836	CTGAGCTTCCCGAAGAGTGCAGCCCGACTTCTGCTGCTGCTGCTGCTGCTGCTGCTG	1895
QY	1683	CTGTACCTGAGAGACTTGCCACCCGACTTCTGATGATGATGATGATGATGATGATGATG	1742
Db	1896	CTGTACCTGAGAGACTTGCCACCCGACTTCTGATGATGATGATGATGATGATGATGATG	1955
QY	1743	ATCATGCTCTCCCTGATGAGAAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	1802
Db	1956	ATCATGCTCTCCCTGATGAGAAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	2015
QY	1803	CTATTTTCAAGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1862
Db	2016	CTATTTTCAAGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2075
QY	1863	AACAGCTGCTGCTACCTTCTCACCTCTTCCACTCTTCTCACTGATGATGATGATGATG	1922
Db	2076	AACAGCTGCTGCTACCTTCTCACCTCTTCCACTCTTCTCACTGATGATGATGATGATG	2135
QY	1923	CAAAAGACCGGGAGCTGCGGGGGAACCTGGGAGTACGTAGTTGCTTTTGGCGTACCA	1982
Db	2136	CAAAAGACCGGGAGCTGCGGGGGAACCTGGGAGTACGTAGTTGCTTTTGGCGTACCA	2195
QY	1983	GAGAAGGCTGTATTAACAAACACAGAGATGCAAGGCTTTTAGAGAAATGTGTTCAA	2042
Db	2196	GAGAAGGCTGTATTAACAAACACAGAGATGCAAGGCTTTTAGAGAAATGTGTTCAA	2255
QY	2043	AACCATGCGTGTATTTTCAACATTAAGAAAGTTTCACTGCTTAAATTTGATTAAC	2102
Db	2256	AACCATGCGTGTATTTTCAACATTAAGAAAGTTTCACTGCTTAAATTTGATTAAC	2315
QY	2103	GGTTTAATCTGCTGCTTCTTCAATTTTGAATTTTAAATAATGCTAGAAATTCCTTC	2162
Db	2316	GGTTTAATCTGCTGCTTCTTCAATTTTGAATTTTAAATAATGCTAGAAATTCCTTC	2375
QY	2163	GAAAGGCTTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2222
Db	2376	GAAAGGCTTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2435
QY	2223	AGCCAGTGTCTTCTTGAAGACCCCTTAATCTTCTTGAATTTTACCAATTT	2282

OY		2103	GTTTAAATTCGCTGTGGTCAATTTTGAGTATTATTTAAAAAATATGCGTAGAATTCCTC	2162
Dd		2120	GTTTTAATTCCTGCTCTGTTCAATTTTGACATATTTTAAAAAATATGCTGATGAATTCCTC	2179
OY		2163	GAAGGCCCTTCAGACACAATGCTATGTTCTGTTCCCAAACCCAGTCTCTCATATTT	2222
Dd		2180	GAAGGCCCTTCAGACACAATGCTATGTTCTGTTCCCAAACCCAGTCTCTCATATTT	2239
OY		2223	AGCCAGTGTCTTTCTTTGAGGACCCCTTAATCTTGCTTTTGAATTTTACCAGTT	2282
Dd		2240	AGCCAGTGTCTTTCTTTGAGGACCCCTTAATCTTGCTTTTGAATTTTACCAGTT	2299
OY		2283	GGATTGGAATGCAGAGGTCTCCAACACTGATTAATAATTTGAAGA	2328
Dd		2300	GGATTGGAATGCAGAGGTCTCCAACACTGATTAATAATTTGAAGA	2345
 RESULT 3 US-09-836-561-2				
	:	:	: Sequence 2, Application US/09836561	
	:	:	: Patent No. US20020038006A1	
	:	:	: GENERAL INFORMATION:	
	:	:	: APPLICANT: Bandman, Olga	
	:	:	: Corley, Neil C.	
	:	:	: Guegler, Karl J.	
	:	:	: TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS	
	:	:	: NUMBER OF SEQUENCES: 6	
	:	:	: CORRESPONDENCE ADDRESS:	
	:	:	: ADDRESSEE: Incyte Pharmaceuticals, Inc.	
	:	:	: STREET: 3174 Porter Drive	
	:	:	: CITY: Palo Alto	
	:	:	: STATE: CA	
	:	:	: COUNTRY: USA	
	:	:	: ZIP: 94304	
	:	:	: COMPUTER READABLE FORM:	
	:	:	: MEDIUM TYPE: Diskette	
	:	:	: COMPUTER: IBM Compatible	
	:	:	: OPERATING SYSTEM: DOS	
	:	:	: SOFTWARE: FastSeq for Windows Version 2.0	
	:	:	: CURRENT APPLICATION DATA:	
	:	:	: APPLICATION NUMBER: US/09/836,561	
	:	:	: FILING DATE: 16-Apr-2001	
	:	:	: CLASSIFICATION: <Unknown>	
	:	:	: PRIOR APPLICATION DATA:	
	:	:	: APPLICATION NUMBER: 09/212,168	
	:	:	: FILING DATE: <Unknown>	
	:	:	: ATTORNEY/AGENT INFORMATION:	
	:	:	: NAME: Billings, Lucy J.	
	:	:	: REGISTRATION NUMBER: 36,749	
	:	:	: REFERENCE/DOCKET NUMBER: PF-0333 US	
	:	:	: TELECOMMUNICATION INFORMATION:	
	:	:	: TELEPHONE: 415-855-0555	
	:	:	: TELEFAX: 415-845-4166	
	:	:	: TELEX: <Unknown>	
	:	:	: INFORMATION FOR SEQ ID NO: 2:	
	:	:	: SEQUENCE CHARACTERISTICS:	
	:	:	: LENGTH: 2550 base pairs	
	:	:	: TYPE: nucleic acid	
	:	:	: STRANDEDNESS: single	
	:	:	: TOPOLOGY: linear	
	:	:	: IMMEDIATE SOURCE:	
	:	:	: LIBRARY: CORNNOT01	
	:	:	: CLONE: 45517	
	:	:	: SEQUENCE DESCRIPTION: SEQ ID NO: 2:	
	:	:	: US-09-836-561-2	
	:	:	: Query Match	
	:	:	: Best Local Similarity 99.8%; Score 2324.4; DB 10; Length 2550;	
	:	:	: Matches 2325; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
OY	:	:	: 3 CCGGCGCTCCCCGGTGTCTCTTCACAGACTGCGTGGCCCTCTGTGAATAAACACCC	62

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2362 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-041-016-1

Query Match 99.9%; Score 2326; DB 9; Length 2362;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3 CCGGGCGCTCCCGGCTGCTCCCTCCAGACGCTCGGCGCCCTCGAATAAACC 62
20 CCGGGCGCTCCCGGCTGCTCCCTCCAGACGCTCGGCGCCCTCGAATAAACC 79
63 GCGAGCCCGAGGCGCCAGAGAGGCGACGTGCGGAGCTCCGCGGGTCCCGCC 122
80 GCGAGCCCGAGGCGCCAGAGAGGCGACGTGCGGAGCTCCGCGGGTCCCGCC 139
123 CGAGCTTCTCTGCGCTTCGATCTCTCTCGCGCGCTTGAGATGCCAGATATA 182
140 CGAGCTTCTCTGCGCTTCGATCTCTCTCGCGCGCTTGAGATGCCAGATATA 199
183 AAGGATCTCACTGTTACATTTCTGCTCTGCTCCAGCCCTGGGAATGACAGGC 242
200 AAGGATCTCACTGTTACATTTCTGCTCTGCTCCAGCCCTGGGAATGACAGGC 259
243 ACAGTGCAGAAAGGCTTGAACCTGATGCGCAGTCAAGACAGTGTATGATATGA 302
260 ACAGTGCAGAAAGGCTTGAACCTGATGCGCAGTCAAGACAGTGTATGATATGA 319
303 ATGCGCAACCATCCCGGCGCTCGCGAGAGACATGATGTGTATACCAAAATGGCG 362
320 ATGCGCAACCATCCCGGCGCTCGCGAGAGACATGATGTGTATACCAAAATGGCG 379
363 GTATTTATGCAATTCCTCCGAGAAACCTGTGTATGAGAGGCGCTTCAACCCCTACTC 422
380 GTATTTATGCAATTCCTCCGAGAAACCTGTGTATGAGAGGCGCTTCAACCCCTACTC 439
423 GACCCCTACTGAGTCCGTACACAGACAGCTGCCACCACTCTCAGTCCAAATATTC 482
440 GACCCCTACTGAGTCCGTACACAGACAGCTGCCACCACTCTCAGTCCAAATATTC 499
483 CAGATCTCCAGGCGCTTATATGCGGCTTGTGATACAGATGATGAAGAACAATG 542
500 CAGATCTCCAGGCGCTTATATGCGGCTTGTGATACAGATGATGAAGAACAATG 559
543 TGTGATGTGAGACAGTGTGCAACAGATTCACCAAGTGCACACCCAGATCTGCAT 602
560 TGTGATGTGAGACAGTGTGCAACAGATTCACCAAGTGCACACCCAGATCTGCAT 619
603 CAATAGTGAAGGCGGTACACCTGCTCTGACCGACGAGATATGCTTGTGAAGGCA 662
620 CAATAGTGAAGGCGGTACACCTGCTCTGACCGACGAGATATGCTTGTGAAGGCA 679
663 GTGCTTGAACATGATGATGCTATGCTTACGCGCAGAGCTCTGCGCAATGTTC 722
680 GTGCTTGAACATGATGATGCTATGCTTACGCGCAGAGCTCTGCGCAATGTTC 739
723 TGAATCTATCTGTACATGCAACCTGTGTTTACCTCAATGAGATGGAAGTCTTG 782
740 TGAATCTATCTGTACATGCAACCTGTGTTTACCTCAATGAGATGGAAGTCTTG 799
783 CCAAGATGTGAACAGTGTGCGACCGAGAAACCTGCGTCAAACTGTGCTCAACCTA 842
800 CCAAGATGTGAACAGTGTGCGACCGAGAAACCTGCGTCAAACTGTGCTCAACCTA 859
843 CGGCTCTTCAATGCGCGCTGTGACCGAGATATGAACTTGAAGAGATGGCGTTCAATG 902
860 CGGCTCTTCAATGCGCGCTGTGACCGAGATATGAACTTGAAGAGATGGCGTTCAATG 919

903 CAGTATATGAGACGAGTACAGCTTCTGTAGTTCCTCTGCAACATGAGTGTGAACA 962
920 CAGTATATGAGACGAGTACAGCTTCTGTAGTTCCTCTGCAACATGAGTGTGAACA 979
963 GCCGGGACATATCTTCTGCTCCGCTTCACAGCTACATCTGTGATGACAAACGAAG 1022
980 GCCGGGACATATCTTCTGCTCCGCTTCACAGCTACATCTGTGATGACAAACGAAG 1039
1023 CTGCCAAGACATCAACGATGTGACACAGAAACCAACAGTGCACACCTCAGACAGCTG 1082
1040 CTGCCAAGACATCAACGATGTGACACAGAAACCAACAGTGCACACCTCAGACAGCTG 1099
1083 CTCAATTTACAAAGGCGCTTCAATGATGACACCCATCCCTGTGAGAGAGCTTATCT 1142
1100 CTCAATTTACAAAGGCGCTTCAATGATGACACCCATCCCTGTGAGAGAGCTTATCT 1159
1143 GAGGATCAGTATAACCGCTGTATGTCTGCTGAGAAACCTGCTGACAGACAGCC 1202
1160 GAGGATCAGTATAACCGCTGTATGTCTGCTGAGAAACCTGCTGACAGACAGCC 1219
1203 CTTTACCATCTTGAACCGGAGCATGAGAGTGTGACAGAGCTCCGCTGACAT 1262
1220 CTTTACCATCTTGAACCGGAGCATGAGAGTGTGACAGAGCTCCGCTGACAT 1279
1263 CTTTCAATGCAAGCAGACCGCTTACCTGCGGCTTATACATTTTCCAGATCAATC 1322
1280 CTTTCAATGCAAGCAGACCGCTTACCTGCGGCTTATACATTTTCCAGATCAATC 1339
1323 TGGAAATGAGGCGCAGAGATTTTACATGCGGCAAAAGGCGCCCATATGACACCTGGT 1382
1340 TGGAAATGAGGCGCAGAGATTTTACATGCGGCAAAAGGCGCCCATATGACACCTGGT 1399
1383 GATGACACGCGCCCATCAAGGCGCCCGGAAATCCAGCTGAGAGCTTGAATGATCACTGT 1442
1400 GATGACACGCGCCCATCAAGGCGCCCGGAAATCCAGCTGAGAGCTTGAATGATCACTGT 1459
1443 CAACACTGTCAATCACTTCAAGAGCAGCTCCGTGATCCGACTGCGGATATATGTGCGA 1502
1460 CAACACTGTCAATCACTTCAAGAGCAGCTCCGTGATCCGACTGCGGATATATGTGCGA 1519
1503 GTACCATCTGAGGCTGCGGCTGAGAGCTCCGAGCTCCCTCTCATTTGGACCAAGGGA 1562
1520 GTACCATCTGAGGCTGCGGCTGAGAGCTCCGAGCTCCCTCTCATTTGGACCAAGGGA 1579
1563 CAGGAAGAGAGGAAATACAGAGATGAGAGACAGACAGACTTAGGCATTTCTCTG 1622
1580 CAGGAAGAGAGGAAATACAGAGATGAGAGACAGACAGACTTAGGCATTTCTCTG 1639
1623 CTGAACGTTTCCCGAAGAGTACGCCCGACTTCTGACTCTCAGCTGTACTATTTGACGA 1682
1640 CTGAACGTTTCCCGAAGAGTACGCCCGACTTCTGACTCTCAGCTGTACTATTTGACGA 1699
1683 CCTGTACCTCTGACAGGACTTGCACACCCAGTTCCTATATACAGTATCAAAATATTT 1742
1700 CCTGTACCTCTGACAGGACTTGCACACCCAGTTCCTATATACAGTATCAAAATATTT 1759
1743 ATCAATGCTCCCGATGAGATTTGTGTAAATTTTCAAGGCTTCAAGTTATTTTCA 1802
1760 ATCAATGCTCCCGATGAGATTTGTGTAAATTTTCAAGGCTTCAAGTTATTTTCA 1819
1803 CTATTTTCAAAAGAAATAGATTAGTTTGCAGGAGCTGAGTCTATGTTCAAAAGCTGTG 1862
1820 CTATTTTCAAAAGAAATAGATTAGTTTGCAGGAGCTGAGTCTATGTTCAAAAGCTGTG 1879
1863 AACAGCTTGTGTACCTTCTTACCTCTTCACTCTCTTCACTGTGTACTGCTTTG 1922
1880 AACAGCTTGTGTACCTTCTTACCTCTTCACTCTCTTCACTGTGTACTGCTTTG 1939
1923 CAAAGACCGGAGAGTGTGCGGAGAACCTGAGAGTATGCTTGTGTTTGTGCTACACA 1982
1940 CAAAGACCGGAGAGTGTGCGGAGAACCTGAGAGTATGCTTGTGTTTGTGCTACACA 1999

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 11:41:03 ; Search time 379.284 Seconds

(Without alignments)
9581.191 Million cell updates/sec

Title: US-09-674-379A-12

Perfect score: 2328
Sequence: 1 gaccgcgcgcctccccctg.....tgatttaataattgagaga 2328

Scoring table: IDENTITY-NUC
Gapop 10.0 , Gapext 1.0

Searched: 1085931 seqs, 780495707 residues

Total number of hits satisfying chosen parameters: 2171862

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCR_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCRUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2326	99.9	2362	9	US-10-041-016-1 Sequence 1, Appl 1
2	2326	99.9	2362	10	US-09-083-002-1 Sequence 1, Appl 1
3	2324.4	99.8	2550	10	US-09-836-561-2 Sequence 2, Appl 1
4	2322.8	99.8	2609	9	US-10-066-500-14 Sequence 14, Appl 1
5	2322.8	99.8	2609	9	US-10-174-580-407 Sequence 407, App
6	2322.8	99.8	2609	9	US-10-176-758-407 Sequence 407, App
7	2322.8	99.8	2609	9	US-10-175-737-407 Sequence 407, App
8	2322.8	99.8	2609	9	US-10-173-706-407 Sequence 407, App
9	2322.8	99.8	2609	9	US-10-175-738-407 Sequence 407, App
10	2322.8	99.8	2609	9	US-10-175-752-407 Sequence 407, App
11	2322.8	99.8	2609	9	US-10-176-482-407 Sequence 407, App
12	2322.8	99.8	2609	9	US-10-176-757-407 Sequence 407, App
13	2322.8	99.8	2609	9	US-10-176-913-407 Sequence 407, App
14	2322.8	99.8	2609	9	US-10-180-552-407 Sequence 407, App
15	2322.8	99.8	2609	9	US-10-180-557-407 Sequence 407, App
16	2322.8	99.8	2609	9	US-10-173-700-407 Sequence 407, App
17	2322.8	99.8	2609	9	US-10-174-572-407 Sequence 407, App
18	2322.8	99.8	2609	9	US-10-174-579-407 Sequence 407, App
19	2322.8	99.8	2609	9	US-10-174-582-407 Sequence 407, App

20 2322.8 99.8 2609 9 US-10-174-588-407 Sequence 407, App
21 2322.8 99.8 2609 9 US-10-175-739-407 Sequence 407, App
22 2322.8 99.8 2609 9 US-10-175-740-407 Sequence 407, App
23 2322.8 99.8 2609 9 US-10-175-743-407 Sequence 407, App
24 2322.8 99.8 2609 9 US-10-176-488-407 Sequence 407, App
25 2322.8 99.8 2609 9 US-10-176-492-407 Sequence 407, App
26 2322.8 99.8 2609 9 US-10-176-474-407 Sequence 407, App
27 2322.8 99.8 2609 9 US-10-176-750-407 Sequence 407, App
28 2322.8 99.8 2609 9 US-10-176-985-407 Sequence 407, App
29 2322.8 99.8 2609 9 US-10-176-987-407 Sequence 407, App
30 2322.8 99.8 2609 9 US-10-176-991-407 Sequence 407, App
31 2322.8 99.8 2609 9 US-10-176-992-407 Sequence 407, App
32 2322.8 99.8 2609 9 US-10-176-993-407 Sequence 407, App
33 2322.8 99.8 2609 9 US-10-184-658-407 Sequence 14, Appl 1
34 2322.8 99.8 2609 9 US-10-002-796-14 Sequence 14, Appl 1
35 2322.8 99.8 2609 9 US-10-066-494-14 Sequence 14, Appl 1
36 2322.8 99.8 2609 9 US-10-173-695-407 Sequence 407, App
37 2322.8 99.8 2609 9 US-10-173-697-407 Sequence 407, App
38 2322.8 99.8 2609 9 US-10-173-705-407 Sequence 407, App
39 2322.8 99.8 2609 9 US-10-173-705-407 Sequence 407, App
40 2322.8 99.8 2609 9 US-10-174-576-407 Sequence 407, App
41 2322.8 99.8 2609 9 US-10-174-585-407 Sequence 407, App
42 2322.8 99.8 2609 9 US-10-174-586-407 Sequence 407, App
43 2322.8 99.8 2609 9 US-10-175-747-407 Sequence 407, App
44 2322.8 99.8 2609 9 US-10-176-481-407 Sequence 407, App
45 2322.8 99.8 2609 9 US-10-176-485-407 Sequence 407, App

ALIGNMENTS

RESULT 1
US-10-041-016-1
Sequence 1, Application US/10041016
Patent No. US20020165151A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
McCoy, John M.
Racie, Lisa A.
LaValle, Edward R.
Merberg, David
Treacy, Maurice
Evans, Cheryl
Agostino, Michael
Lu, Zhijian
Honjo, Tasuku
TITLE OF INVENTION: SECRETED PROTEINS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/041,016
FILING DATE: 07-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/083,002
FILING DATE: 21-MAR-1998
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: P-41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851

Db 3181 CGAGATGTGACGAGTGTGCGCCGAGCCGCTGCTGCCCGAGGGGCTTGCCTCAACACG 3240
QY 841 TACGGCTCTTTCATCTGCGG---CTGTGACCCAGATATGAATCTTGAGAGATGGCGTT 897
Db 3241 GAGGGCTCTTACCTGCTGACGCTGTGACAGCGGGTACTGGGTGAACGAGATGGCAGT 3300
QY 898 CATTGCACTATATGAGAGATGACGCTTCTGAGTTCCTGCTGCCAATGATGTGTG 957
Db 3301 GCGCTGAAGACTGTGATGAATGTGCTTCCCTGAGTGTGCCCGACAGGCGTCTGACC 3360
QY 958 AACGAGCCGCGACATCTTGTGCTGCTGCTCCAGGCTTACATCTGCTGATGACAC 1017
Db 3361 AATGCTGAGCTCTTCTCTGCAAGAGACTGTGACAGGGCTACCGGCCAACCCCGT 3420
QY 1018 CGAAGTGTGCAAGACATCAACAGATGTGAGACACAGAACACAGTGTGACAGT 1077
Db 3421 GCGAACAGATGCGAAGATGTGATGATGTGAGTGTGCGCCAAAGCAGCTGCGGGAGGC 3480
Y 1078 ACCTGCTACATTTTCAAGGGGGCTTCAAAATGATGACGCCCATCCGCTGTGAGAGCT 1137
Db 3481 GAATGCAAGAACACAGAGAGTGTGCTTACCAATGCTCTGTGACAGGCGTTCAGCTGTG 3540
QY 1138 TATCTGAGGATCAGTATACCGCTGTATGTGTCTGCTGAGAACCTGG 1187
Db 3541 AATGGACCATGTGTGAGGAGCGTGAATGATGTGTGGGGAAGACATTG 3590

RESULT 14
PCT-US95-02251-17
Sequence 17, Application PC/TUS9502251
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING BONE
NUMBER OF INVENTION: CELLS
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: PatentIn Release #1.0, Version
SOFTWARE: #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02251
FILING DATE: CONCURRENTLY HEREMITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/316,650
FILING DATE: 30-SEP-1994
CLASSIFICATION:
APPLICATION NUMBER: US 08/199,780
FILING DATE: 18-FEB-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UMIC009P--
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 5502 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "DNA"
FEATURE:
NAME/KEY: CDS
LOCATION: 1..5502
PCT-US95-02251-17

Query Match
Best Local Similarity 47.4%; Pred. No. 5e-05;
Matches 223; Conservative 0; Mismatches 241; Indels 6; Gaps 2;

2.3%; Score 52.4; DB 5; Length 5502;

QY 724 GGAATCTATTTCTGTACATGACAACTGTTTACCTCAATGATGAGATGAGTCTTGC 783
Db 3121 GGTCTCTTATGATGCTCTGTGAGCGGGCTATGAGTCAACCCAGCAAGAGGGCTGC 3180
QY 784 CAAGATGTGACGAGTGTGCGCCAGCGAACCCTG---CGTGAACCTGCTCAACAC 840
Db 3181 CGAGATGTGACGAGTGTGCGCCAGCGAACCCTGCGCCACAGGGGCTCTGCTCAACAG 3240
QY 841 TACGCTCTTTCATCTGCGG---CTGTGACCCAGATATGAATCTTGAAGAGATGGCGTT 897
Db 3241 GAGGGCTCTTACCTGCTGACAGCTGTGAGAGCGGGTACTGGGTGAACGAAATGGGACT 3300
QY 898 CATTGCACTATATGAGAGATGACGCTTCTGAGTTCCTGCTGCCAATGATGTGTG 957
Db 3301 GCGCTGAAGACTGTGATGAATGTGCTTCCCTGAGTGTGCCCGACAGGCGTGTGACC 3360
QY 958 AACGAGCCGCGACATCTTGTGCTGCTGCTCCAGGCTTACATCTGCTGATGACAC 1017
Db 3361 AATGCTGAGCTCTTCTCTGCAAGAGACTGTGACAGGGCTACCGGCCAACCCCGT 3420
QY 1018 CGAAGTGTGCAAGACATCAACAGATGTGAGACACAGAACACAGTGTGACAGT 1077
Db 3421 GCGAACAGATGCGAAGATGTGATGATGTGAGTGTGCGCCAAAGCAGCTGCGGGAGGC 3480
QY 1078 ACCTGCTACATTTTCAAGGGGGCTTCAATGATGATGACACCCATCCGCTGTGAGAGCT 1137
Db 3481 GAATGCAAGAACACAGAGAGTGTGCTTACCAATGCTCTGTGACAGGCGTTCAGCTGTG 3540
QY 1138 TATCTGAGGATCAGTATACCGCTGTATGTGTCTGCTGAGAACCTGG 1187
Db 3541 AATGGACCATGTGTGAGGAGCGTGAATGATGTGTGGGGAAGACATTG 3590

RESULT 15
PCT-US95-02251-2
Sequence 2, Application PC/TUS9502251
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING BONE
NUMBER OF INVENTION: CELLS
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: PatentIn Release #1.0, Version
SOFTWARE: #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02251
FILING DATE: CONCURRENTLY HEREMITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/316,650
FILING DATE: 30-SEP-1994
CLASSIFICATION:

Db 940 ACAATGA---ATGCCGGGAGAGTAAGTAATGTTGGAAATTTATCATATGGCGGCTTCCGTTGT 996

Qy 1111 ATCGACCCCATCCGGCTGTGAGGACCCTTATCTGAGATCAAGTATTAACCCGCTGTATGTGT 1170

Db 997 TATTCACGAATATCCTTGTCAAGATCCCTACTATCTTAACACAGAGAACCAGATGTGTTTGC 1056

Qy 1171 CTTCTCTAGAAACCTGGCTGACAGAGCACGCCCTTTACACTCTTGTATCCGGACATGGAC 1230

Db 1057 CCAATCTCAATATGCATATGCGCGAAGATGCCCCAGTCAATATAGTCTCAATATCATAGNC 1116

Qy 1231 GTGGTGTCAAGAGCCTCGTTCGCCGACACTCTTCAATATGCACAGACAGACCCGGTAC 1290

Db 1117 ATCCGATCTGTATAGTGTGTGTCATCAGACACTCTTCCAGATACAGGCGCACACATTTAT 1176

Qy 1291 CCTGGGGCCTATTTACATTTTCCAAATCAAAATCTGGAGTGAGGCGACAGATTTTACATG 1350

Db 1177 GCCAACACCAATCACTACTTTTCGATTTAAATCTTGAAATGGAATGAGAGATTTCTACCTA 1236

Qy 1351 CGGCACAAAGGGCCCCCATCACTAGTCCACCCCTGGTATGACAGGCCCATCAAGAGGCCCGG 1410

Db 1237 CGACAAACCAAGTCCCTGTAGTGAACAATCCTGTGCTCGGAAGTCATATTATCAGACCAAGA 1296

Qy 1411 GAAATCCAGCTGGACTTGGAAATATATCACTGTCAACACTGTCAATCAACTTCAGAGGACG 1470

Db 1297 GAACATATTCGTGACACCTGGAGATCTCTACAGCTCAGCAGTATAGGACCTTCCGACAAAGC 1356

Qy 1471 TCCGTATCCGACTGGGATATATGTATGCGAGTACCACTTCT 1513

Db 1357 TCTGTGTTAAGATTGACATATATATGAGGGCCATTTTCATTTT 1399

RESULT 8
 US-08-897-443-2
 : Sequence 2, Application US/08897443
 : Patent No. 5981263
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Hillman, Jennifer L.
 : APPLICANT: Lal, Preeti
 : APPLICANT: Corley, Neil C.
 : APPLICANT: Shah, Purvi
 : APPLICANT: Kaser, Mathew
 : TITLE OF INVENTION: HUMAN MATRILIN-3
 : NUMBER OF SEQUENCES: 4
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Incyte Pharmaceuticals, Inc.
 : STREET: 3174 Porter Drive
 : CITY: Palo Alto
 : STATE: CA
 : COUNTRY: USA
 : ZIP: 94304
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: DOS
 : SOFTWARE: FastSeq for Windows Version 2.0
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/897,443
 : FILING DATE: Filed Herewith
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER:
 : FILING DATE:
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Billings, Lucy J.
 : REGISTRATION NUMBER: 36,749
 : REFERENCE/DOCKET NUMBER: PP-0348 US
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 415-855-0555
 : TELEFAX: 415-845-4166
 : INFORMATION FOR SEQ ID NO: 2:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 3373 base pairs
 : TYPE: nucleic acid

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: UTRSNCT02
; CLONE: 681719
;
US-08-897-443-2

```

Query Match	Score	DB 2;	Length
4.68;	106.8;		3373;

Matches 211; Conservative 0; Mismatches 152; Indels 3; Gaps 1;

695 ACTGCCAGCAGCTCTGTGGGATGTTCCCTGATCCTATTCTTGACATGCACCCCTGTT 754

920 ACTGTGAGCAGCTCTGTGTGAATGTGCCGGGCTCCTTCGTCTGCCAGTGCTACAGTGGCT 979

755 TTACCCCTCATGAGGATGGAAGGTCTTGCCACAGATGTGAACGAGTGTGCCACCGAGAACC 814

980 ACGCCCTGGCTGAGGATGGGAAGAGCTGTGTGGCTGTGGACTACTGTGCCCTCAGAAACC 1031

815 CC--TCGCTGCAACCTGCGTCAACACCTACGGCTCTTTCATCTGCCGCTGACCCAG 871

1040 ACGGATGTGACATGAGTGTAAATGCTGATGGCTCCCTACCTTTGGCAGTGCATGAG 1095

8/2 GAATATGACCTTGGAGAGATGGCGTTCATTGCAAGTGAATATGACGAGTGCAGCTTCTCTG 931

1100 GATTTGGCTCTTAAACCAGATMAAAACGTCACAAAGATAGACTACTGTGCTCATCTA 1150

532 AGGAGCCTGGACATGAGGAGGAGGACCGCCCGGCACCACTTCCCTCCCTCC

1100 ATCTGGGATGTCAGCAGCGAGTGTATTTACATCAAGATGATTCCATAACCCTCCGCATCCCATTAAT

[illegible][illegible][illegible]

RESULT 9

6 IN

Sequence 278, Application US/09643597

GENERAL INFORMATION:

APPLICANT: Fan, Liqun

APPLICANT: Bangur, Chaitanya S.

APPLICANT: Fanger, Gary R.

APPLICANT: Wang, Aijun
APPLICANT: Skolky, Yassir A W

APPLICANT: Henderson, Robert A.
APPLICANT: McNeill, Patricia D.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

FILE REFERENCE: 210121.455C11
CURRENT APPLICATION NUMBER: IIS/09/643.597

CURRENT FILING DATE: 2000-08-21
NUMBER OF SEO ID NOS: 369

SOFTWARE: FASTSEQ IOT WINDOWS VERSION 3.0
SEO ID NO 278

LENGTH: 401
TYPE: DNA

ORGANISM: homo sapien
FEATURE:

NAME/KEY: MISS_CARLUE
LOCATION: (1) ... (401)

IS-09-643-597-278

```

; Sequence 259, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011C1
; CURRENT APPLICATION NUMBER: US/09/188.930A
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 259
; LENGTH: 1018
; TYPE: DNA
; ORGANISM: Human
US-09-188-930-259

```

```

Query Match      14.2%; Score 331.4; DB 3; Length 1018;
Best Local Similarity 62.4%; Pred. No. 1.9e-87;
Matches 537; Conservative 0; Mismatches 321; Indels 3; Gaps 1;

```

```

OY 686 GCTATGTTACTGCGACAGCTGTGCGAATGTTCTGATCTCTTCTGTAACGCA 745
DB 1 GCTACCGCTACTGCGACAGCTGTGCGAATGTTCTGATCTCTTCTGTAACGCA 60
OY 746 ACCCTGTTTACCTCAATGAGATGAGAGTCTTCTGCAAGATGTAAGAGTGTGCA 805
DB 61 AGCCGGGCTTCCAGTGGGGCCCTCAACAGCGCTCTGTTGTGATGTAAGAGTGTGCA 120
OY 806 CCGAAGACCCCTGCGCGCAACCTGCTCAACACCTTACGCGCTCTTATCTGCGCTGTG 865
DB 121 TGGGGGCCCCATGCGAGCGCGCTGCTCAACCTCAATGAGAGCTTCTGTCGTGCGC 180
OY 866 ACCAGATATGATCTGAGAGATGAGAGTGTGCAAGTGTGTAAGAGTGTGCA 925
DB 181 ACCAGGATATGATCTGAGAGTGTGCAAGTGTGTAAGAGTGTGTAAGAGTGTGCA 240
OY 926 TCTGTGATCTCTGCGCAACATGATGTGTGTAAGAGTGTGTAAGAGTGTGCA 985
DB 241 ACTCAGCTACTCTGTCTGATGACCGCTGCTGTAAGAGTGTGTAAGAGTGTGCA 300
OY 986 GCCCTCAGAGCTATCTGCTGATGATGATGATGATGATGATGATGATGATGATG 1045
DB 301 GCCCAGAGGTTTACCACTGTG--GCCACAGCGCTGTGCAAGATGATGATGATG 357
OY 1046 AGCAGAGAGACCAAGCTGCAACCTGCAAGAGTGTGCAAGTGTGCAAGAGGCTTCA 1105
DB 358 AGTGTGTGCGGACCAAGTGTGCGAGGCGCAACCTGCTGCAAGTGTGCAAGAGGCTTCA 417
OY 1106 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1165
DB 418 GCTGTGTGAGACCAACAGCTGCTGAGAGCGCTGCAAGTGTGCAAGAGGCTTCA 477
OY 1166 TGTGTCTGCTGAGAGACCTGTGCTGAGAGACCGAGCTTACCATCTTACCGGACA 1225
DB 478 TCTGTGCGGCTTCCAACTCTGATGTGAGAGAGCGCTTACCATCTTACCGGACA 537
OY 1226 TGGAGTGTGTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1285
DB 538 TGACCATCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 597
OY 1286 GCTACCTGCGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1345
DB 598 TCTACCTGCGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 657
OY 1346 ACATGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1405
DB 658 ACATGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 717

```

```

OY 1406 CCGGGAATTCAGCTGACTTGGAAATGATCACTGTCAACACTGTCAATCAAG 1465
DB 718 CCGGGAATTCAGCTGACTTGGAAATGATCACTGTCAACACTGTCAATCAAG 777
OY 1466 GCAGTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1525
DB 778 CAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 837
OY 1526 GAGAGCTCGAGAGCTGCTCT 1546
DB 838 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 858

```

```

RESULT 7
US-09-248-757-1
; Sequence 1, Application US/09248757
; Patent No. 6417342
; GENERAL INFORMATION:
; APPLICANT: STONE, EDWIN M.
; APPLICANT: SHEPHERD, VAL C.
; TITLE OF INVENTION: MACULAR DEGENERATION DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: UIA-018.02
; CURRENT APPLICATION NUMBER: US/09/248.757
; CURRENT FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2512
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-248-757-1

```

```

Query Match      10.4%; Score 241.4; DB 4; Length 2512;
Best Local Similarity 54.6%; Pred. No. 9e-61;
Matches 548; Conservative 0; Mismatches 446; Indels 9; Gaps 3;

```

```

OY 514 GATATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 573
DB 403 GCTTACGAGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 462
OY 574 CACAGTGCACCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 633
DB 463 CACAGTGCACCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 522
OY 634 ACCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 690
DB 523 CCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 582
OY 691 GGTACTGCGACAGCTGTGAGATGATGATGATGATGATGATGATGATGATGATG 750
DB 583 CCAATATGCGACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 642
OY 751 GGTATACCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 810
DB 643 GGTATACCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 702
OY 811 AACCCCTGCTGCAACCTGCTGCAACCTGCTGCAACCTGCTGCAACCTGCTGCAAC 870
DB 703 AATCAATGCTGCAACCTGCTGCAACCTGCTGCAACCTGCTGCAACCTGCTGCAAC 762
OY 871 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 930
DB 763 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 822
OY 931 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 990
DB 823 AGCTACCTGCTGCAACCTGCTGCAACCTGCTGCAACCTGCTGCAACCTGCTGCAAC 882
OY 991 CAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1050
DB 883 CAGGATA--CCAAGTGTGAGAGTGAAGATGATGATGATGATGATGATGATGATG 939
OY 1051 AGAAGCAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1110

```

Db	1226	GTCTACCCCGGTGGCTTCACAAATGCCCTTTTCAGATCCGTGTGGAAACTGCGACAGGGGAGCTTT	1285
Qy	1345	TACATGGGGCAAAACGGGGCCCCCATATAGTGGCCACCTGGTGATGAGACAGCCCCCATTAAGAGG	1404
Db	1286	TACTTTAGGGCAATCAACAACGTATAGGCGCCATGTGGTCTCGCCCGGGCCGGTACGGGC	1345
Qy	1405	CCCCGGGAAATCCACGTGGAGCTTGGAAATGATCACTGTCACACACTGTATCAACATTGAGA	1464
Db	1346	CCCCGGGAGTACAGCGTGGACCTGGAATGATGTCACCAATGAATTTCCCTCATAGCTACCGG	1405
Qy	1465	GGCAGCTCCGTGATCCACATGCGGATATATATGTGCGAGTATCCCATTTGTGAGCCTCGGGC	1524
Db	1406	GCCAGCTCTGTACTAGAGGCTCAACCGCTCTTTGTATAGGGGCTTACACTTCTGAGAGACAGA	1465
Qy	1525	TGAGAGCTCCGACGCTGCTCTT	1546
Db	1466	GGGAGCCACCTCCCTCGACGCT	1487

RESULT 5
US-08-98

Sequence 2, Application US/08980514
Patent No. 6004753

APPLICANT: Yue, Henry
 APPLICANT: Guegler, Karl J.
 APPLICANT: Shan, Puiyi
 TITLE OF INVENTION: HUMAN SI-5-ECMP-LIKE PROT
 TITLE OF INVENTION: EIN
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Dr.
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/980,514
 FILING DATE: Filed Herewith
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0436 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-855-0555
 TELEFAX: 650-845-4166
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2018 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: BRSTN0T13
 CLONE: 2786449
 US-08-980-514-2

Query Match	16.6%;	Score 386;	DB 3;	Length 2018;
Best Local Similarity	61.3%;	Pred. No. 2.5e-103;		
Matches 639;	Conservative 0;	Mismatches 400;	Indels 3;	Gaps 1.

505 TGGCCGCTTTMGGATACAGATGGATGAAAGCAACCAATGTGTGGATGTGGACGAGTGTGCA 564

Db	530	TGCCACCAAGGCTATGAGCCCGGAGATCAGGACAGCTGTGTGATGTGGACGAGTGTGCC	589
Qy	555	ACAGATTCCACACCAAGTCGACACCCACCCAGATCTGCAATCAATCTAGAAAGCGGGTACAC	624
Db	590	CAGGCGCCCTGACGAGCTGTGCCCCAGCCAGGACGCTCAATCACTTGGCCGTCTCTATCAG	649
Qy	625	TGTCCTGTACACGAGCGATATTGGCTTCTGGAAAGCCAGTGCCTTAGACATGTAGTAATGT	684
Db	650	TGCACCTGCCCTGATGGTTTACCCGCAAGATCGGGCCGAGGTGTGTGGACATAGACAGTGC	709
Qy	685	CGTATGGTTACTGCCAGACAGCTCTGTGCGAATGTTCTCGATCCATCTTCTGTACATGC	744
Db	710	CGTACACCGCTACTGCCAGACCGCTGCGTGAACCTGCTGGCTCTCTTCCGCTGCCAGTGC	769
Qy	745	AAACCTGGTTTATACCCATCAGAGGATGGAAAGTCTGGCAAGATGTGTAACAGAGTGTCC	804
Db	770	GAGCCGGGGCTTCCAGCTGGGGGCTTAAACACCGCTCTCTGTGTGATGTAAAGAGTGTAC	829
Qy	805	ACCGAAGAACCCCTGGTGTGCAAACTGTGCTCAACACTTACGGCTCTTCTATCTGCCGCTGT	864
Db	830	ATGGGGGGCCCATGTGGAGACAGCGGTGCTTCAACTCTATGTGGACCTTCTGTGTGCCGTC	889
Qy	865	GACCCAGGATATTGAACCTTGAGGAAGATGGCGTTCAATTGCAATGTATGTGACGATGCAC	924
Db	890	CACCAAGGCTATTGAGCTGTGCATCGGAGATGGCTTCTCTGCACTGTATGTGAGAGTGTAGC	949
Qy	925	TTTCTGTAGTTTCTCTGTGCAACATGAGTGTGTGAACAGCCCGGACATATTCTTGCTCC	984
Db	950	TACTCTACGCTACCTCTCTGTCACTACCGCGCTGCGTCAACAGCCAGCCGCTTTCTCTGCCAC	1009
Qy	985	TGCCCTCCAGGCTACATCTCTGATGATGACAACGGAAGCTGCCAAGACATCAAGAAATGT	1044
Db	1010	TGCCACAGGGTTTACAGACTGCTG---GCCACAGCCCTCTGCCAAGACATGTATGATAGTGT	1066
Qy	1045	GAGCAGAGGAACACACAGCTGCACTGTGAGAGAGAGTCTCAATTTTCAAGGGGGCTTC	1104
Db	1067	GAGTCTGTGTGCGACACAGTGTCTCGAAGGCCCAACCTGTGTCACTTCCATAGGGGGCTAC	1128
Qy	1105	AAATGCACTCGACCCCATCCGCTGTGTGAGAGGCTTATCTGAGGATCAGTATTAACCGCTGT	1164
Db	1127	CGCTGCGTGTGAGACCAACACGCGTGGTGAAGCCCTATCAACAGTCTGTGAGAACCGCTGT	1186
Qy	1165	ATGTGTCTGTGTGAAACCTGTGGCTGCAAGAGACACAGCCCTTTTACCATCTTGTACCGGAC	1224
Db	1187	CTGTGCGCGGCTCCACACCTCTATGTGTGAAGAGCGCTTATCATCTGTGTACCGCTTAC	1246
Qy	1225	ATGGAAGTGTGTGAGAGCGCTCGGTCGCCGCTGTGACATCTTCCAAATGCCAACGACGAC	1284
Db	1247	ATGACCATTCACCTGGAGCGGAGCGGTGCGCGCTGTGCGTTCAGATCCAGACGCGACTTC	1306
Qy	1285	CGTCAACCTTGGGGCTTATTACATTTTCCAGATCAAAATGTGGAAATGAGAGCGAGAAATTT	1344
Db	1307	GCTTACCCCGGGTGCCTAACATGCGTTTCAGATTCGTTGTGGAACTCGCAGGGGAGCTTT	1366
Qy	1345	TACATGTGGGCAACGCGGCCCATCATGCTGCCACCTGTGTGTGATGACAGCGCCCATCAAGGG	1404
Db	1367	TACATTTAGGCAAAATCAACACGTCAGGCCCATGTGTGTCTTCGCCCGGCCGTGTACGGGC	1428
Qy	1405	CCCCGGGAAATACAGCTGGAGCTGTGGAAATGTACCTGTCAACACTGTATCAACTTTCGA	1464
Db	1427	CCCCGGGAGTACGTGTCTGGACCTGGAGATGTGTCAACATGAATTTCCCTATAGCTTACC	1486
Qy	1465	GGCAGCTTCCGTATCCGATCGGATATATATGTGTGCGAGTACCCATTCCTGAAGCTCGGGC	1524
Db	1487	GCCAGCTCTGTACTGAGGCTACCGCTTTTGTAGGGGGCTCTACACTTCTGTAGAGACAGGA	1546
Qy	1525	TGAGCCTTCGAGCGCTGCTCT 1546	
Db	1547	GGGAGCCACCTCTCCCTGCACT 1568	

RESULT 6

RESULT 6
US-09-188-930-259


```

Db      1416 CTTTACCATCTGTACCGGAGACATGAGCTGTGTGTCAGAGACGCTCCGTTCCCGGTGACAT 1475
QY      1263 CTTCCAAATGCAAGCCAGCAGCCGCTACCTGGGGCCTATTACATTTTCCAGATAAATC 1322
Db      1476 CTTTCAAAATGCAAGCCAGCAGCCGCTACCTGGGGCCTATTACATTTTCCAGATAAATC 1535
QY      1323 TGGGAATGAGGAGAGAAATTTTACATGCGGCAAAACGGGCCCTACCTGCGACCCCTGGT 1382
Db      1536 TGGGAATGAGGAGAGAAATTTTACATGCGGCAAAACGGGCCCTACCTGCGACCCCTGGT 1595
QY      1383 GATGACAGCGCCCATCAAGGGCCCGGGAAATCAGCTGAGCTTGAATGATCAGCT 1442
Db      1596 GATGACAGCGCCCATCAAGGGCCCGGGAAATCAGCTGAGCTTGAATGATCAGCT 1655
QY      1443 CAACACTGTCATCACTTCAGAGACACTCCGTCATCCGATGGGGATATGTGTGCA 1502
Db      1656 CAACACTGTCATCACTTCAGAGACACTCCGTCATCCGATGGGGATATGTGTGCA 1715
QY      1503 GTACCCATTCAGAGCTCGGGCTGAGAGCTCCGAGCTGCTCATATGGACCAAGGA 1562
Db      1716 GTACCCATTCAGAGCTCGGGCTGAGAGCTCCGAGCTGCTCATATGGACCAAGGA 1775
QY      1563 CAGGAGAGAGAGAGAAATTAACAGAGAGATGAGAGCAGACAGACAGCTTAGGCAATTCCTG 1622
Db      1776 CAGGAGAGAGAGAGAAATTAACAGAGAGATGAGAGCAGACAGACAGCTTAGGCAATTCCTG 1835
QY      1623 CTGAACGTTTCCCGGAAGAGTCAAGCCCGGACTCCGATCTCACTCACTCACTCACTCACTCA 1682
Db      1836 CTGAACGTTTCCCGGAAGAGTCAAGCCCGGACTCCGATCTCACTCACTCACTCACTCACTCA 1895
QY      1683 CCTGTACCCCTGAGAGACTTGCACACCCAGTTCCTATGATACAGTATTCAAAAGATAT 1742
Db      1896 CCTGTACCCCTGAGAGACTTGCACACCCAGTTCCTATGATACAGTATTCAAAAGATAT 1955
QY      1743 ATCATTTGCTCCCTGATGAGAGATTTGGTGAATTTTCAAGCCCTCAGTTATTTTCA 1802
Db      1956 ATCATTTGCTCCCTGATGAGAGATTTGGTGAATTTTCAAGCCCTCAGTTATTTTCA 2015
QY      1803 CTATTTTCAAGAAATATAGTATGTTGGGGGGGTGAGTCTATGTTTCAAGAGCTGTG 1862
Db      2016 CTATTTTCAAGAAATATAGTATGTTGGGGGGGTGAGTCTATGTTTCAAGAGCTGTG 2075
QY      1863 AACAGCTGCTGTCACCTTCTTCACTCTTCACTCTTCACTGTTACTGCTTGTG 1922
Db      2076 AACAGCTGCTGTCACCTTCTTCACTCTTCACTCTTCACTGTTACTGCTTGTG 2135
QY      1923 CAAAGACCCGGGAGCTGGCGGGGAAACCTGGAGTACGTTTGTTCGTTACCA 1982
Db      2136 CAAAGACCCGGGAGCTGGCGGGGAAACCTGGAGTACGTTTGTTCGTTACCA 2195
QY      1983 GAGAAGGCTATGTAACAAACACAGAGAGATCGAAGGTTTTAGAGATGTGTTCAA 2042
Db      2196 GAGAAGGCTATGTAACAAACACAGAGAGATCGAAGGTTTTAGAGATGTGTTCAA 2255
QY      2043 AACCATGCTGTATTTTCAACATAAAGAGATGTTGCTTAAATTTGTATAC 2102
Db      2256 AACCATGCTGTGTATTTTCAACATAAAGAGATGTTGCTTAAATTTGTATAC 2315
QY      2103 GGTATTAATCTGCTGTGTCATTTTGTAGTATTTTAAATAATGTCTAATTCCTTC 2162
Db      2316 GGTATTAATCTGCTGTGTCATTTTGTAGTATTTTAAATAATGTCTAATTCCTTC 2375
QY      2163 GAAAGGCTTTCAGACATGCTATGTCTCTCCAAACCCAGTCTCTCCATTTT 2222
Db      2376 GAAAGGCTTTCAGACATGCTATGTCTCTCCAAACCCAGTCTCTCCATTTT 2435
QY      2223 AGCCAGTCTTTTCTTTCAGAGACCCCTTAATCTTCTTTCAGATTTTACCCTAAT 2282
Db      2436 AGCCAGTCTTTTCTTTCAGAGACCCCTTAATCTTCTTTCAGATTTTACCCTAAT 2495
QY      2283 GGATTGAGTGCAGAGGTCTCCAAACTGATTAATATTGAAGGA 2328

```

```

Db      2496 GGATTGAGTGCAGAGGTCTCCAAACTGATTAATATTGAAGGA 2541

RESULT 3
US-09-188-930-67
; Sequence 67, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorina
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188.930A
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 67
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Rat
US-09-188-930-67

Query Match      38.5% Score 896; DB 3; Length 1260;
Best Local Similarity 85.1%; Pred. No. 2,3e-253;
Matches 1064; Conservative 5; Mismatches 150; Indels 31; Gaps 5;

QY      506 GCCCTTTGGATACAGATGATGAAAGCAACCAATGTGTGATGTGACAGTGTGCA 565
Db      7 GTGCTTTGGGTATCAATGATGATGAAAGCAACCAATGTGTGATGTGACAGTGTGCA 66
QY      566 CAGATTTCCACAGCTGCAACCCCAAGATCTGATCAATATCTAAGCGGGTACACT 625
Db      67 CAGATTTCCACAGCTGCAACCCCAAGATCTGATCAATATCTAAGCGGGTACACT 126
QY      626 GCTCTGACAGCAGAGATATGCTTCTGAAAGGCAAGTGTGATGATGATGATG 685
Db      127 GCTCTGACAGCAGAGATATGCTTCTGAAAGGCAAGTGTGATGATGATGATG 186
QY      686 GCTATGTTACTGCCAGACCTGTGCGAATGTTCTGATATCTTATCTGTATACATCA 745
Db      187 GCTATGTTACTGCCAGACCTGTGCGAATGTTCTGATATCTTATCTGTATACATCA 246
QY      746 ACCCTGTTTACCTCAATGAGATGAGAGCTTGTCCAGATGTGAACAGTGTGCA 805
Db      247 ACCCTGTTTACCTCAATGAGATGAGAGCTTGTCCAGATGTGAACAGTGTGCA 306
QY      806 CCGAGAACCCCTGCTGCAACCTGCGTCAACACCTACGCTTTCATCTGTGCGCTGTG 865
Db      307 CCGAGAACCCCTGCTGCAACCTGCGTCAACACCTACGCTTTCATCTGTGCGCTGTG 366
QY      866 ACCCAGATATGAACCTTGAGAGAGATGCGTTCAATTCAGATGTGACAGAGTGCAGT 925
Db      367 ACCCAGATATGAACCTTGAGAGAGATGCGTTCAATTCAGATGTGACAGAGTGCAGT 426
QY      926 TCTGAGTCTGCTGCGCAACATGAGTGTGAACCGCGGCAACATCTTCTGCTCT 985
Db      427 TCTGAGTCTGCTGCGCAACATGAGTGTGAACCGCGGCAACATCTTCTGCTCT 486
QY      986 GCCCTCAGGCTCATCTCTGATGATGACCAACGAGCTGCGCAAGATCAACAGATGTG 1045
Db      487 GCCCTCAGGCTCATCTCTGATGATGATGACCAACGAGCTGCGCAAGATCAACAGATGTG 546
QY      1046 AGCAGAGAAACACAGCTGCAACCTGCAAGAGCTGCTATCAATTTACAGGGGCTTCA 1105
Db      547 AGCAGAGAAACACAGCTGCAACCTGCAAGAGCTGCTATCAATTTACAGGGGCTTCA 606
QY      1106 AATGATGAGCCGATCCGCTGTGAGAGAGCTTATCGAGATGATGATGATGATGATG 1165
Db      607 AATGATGAGCCGATCCGCTGTGAGAGAGCTTATCTGATGATGATGATGATGATGATG 666

```


QY	2163	GAAGGCGCTTAGACACACATGCTATGTTCTGTCTTCCAAACCCAGTCTCTCCATTTT	2222
Db	2376	GAAGGCGCTTAGACACACATGCTATGTTCTGTCTTCCAAACCCAGTCTCTCCATTTT	2435
QY	2223	AGCCAGAGTTCTTTTGAGGAGCCCTTAATCTGCTTCTTTTGAATTTTACCCAAT	2282
Db	2436	AGCCAGAGTTCTTTTGAGGAGCCCTTAATCTGCTTCTTTTGAATTTTACCCAAT	2495
QY	2283	GGATTGGAATCAGAGGCTCTCCAACGATTAATATTTGGAAGA	2328
Db	2496	GGATTGGAATCAGAGGCTCTCCAACGATTAATATTTGGAAGA	2541

RESULT 2
US-09-212-168-2
; Sequence 2, Application US/09212168
: Patent No. 6303765

```

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSER: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/212,168
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/884,072
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0333 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2550 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: CORNNOT01
CLONE: 45517
US-09-212-168-2

```

```
Query Match          99.8%; Score 2324.4; DB 4; Length 2550;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2325; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY	Db	QY	Db
3	216	63	276
CCCCGGCCTCTCCCGGTCTCCTCAGACACGGTGGGCCCCCTCMGGAATAAACACC	CCGGGCGCTCCCCGCTCTCTCTCAGACACGCTGGCCCCCTCTGGAAATAAACACC	GGCAGCCCCGAGGGGCCCCAGAGAGGGCCGACGTGTCCTCTCCGGGGGTCCGCCCG	GGCAGCCCCGAGGGGCCCCAGAGAGGGCCGACGTGTCCTCTCCGGGGGTCCGCCCG
62	275	122	335

QY	123	CGAGCTTCTCTTCGCGCTTCGCATCTCCCTCGCGGCTTGTGACATGCGAGAAATAA	102
Db	336	CGAGCTTCTCTTCGCGCTTCGCATCTCCCTCGCGGCTTGTGACATGCGAGAAATAA	399
QY	183	AAGGATACACTGTTACCAATCTGGCTCTGTCTTCCAAAGCCCTGGGAATGCAGAGC	242
Db	396	AAGGATACACTGTTACCAATCTGGCTCTGTCTTCCAAAGCCCTGGGAATGCAGAGC	455
QY	243	ACAGTCCAGAAATGGCTTTGACCTGGATGCGCGTGCAGACATGTTTATGATGATGA	302
Db	456	ACAGTCCAGAAATGGCTTTGACCTGGATGCGCGTGCAGACATGTTTATGATGATGA	515
QY	303	ATGCGGAACCAATCCCGGAGGCTGCGGAGAGACATGATGTGTGTTAAACCAATGGCG	362
Db	516	ATGCGGAACCAATCCCGGAGGCTGCGGAGAGACATGATGTGTGTTAAACCAATGGCG	575
QY	363	GTAATTTATGCAATTCGCCGAGCAAAACCTGTGATGAGAGGCCCTACTCGAACCCCTACTC	422
Db	576	GTAATTTATGCAATTCGCCGAGCAAAACCTGTGATGAGAGGCCCTACTCGAACCCCTACTC	635
QY	423	GACCCCTCACTCAGGTCCGTACCCACAGCACTGCCCCACACTCTCAGCTCCAAACATATCC	482
Db	636	GACCCCTCACTCAGGTCCGTACCCACAGCACTGCCCCACACTCTCAGCTCCAAACATATCC	695
QY	483	CACGATCTCAGGCGCTTATATGCGGCTTTGGATACAGATGATGAAAGCAACCAATG	542
Db	696	CACGATCTCAGGCGCTTATATGCGGCTTTGGATACAGATGATGAAAGCAACCAATG	755
QY	543	TGTGATGTGGAGAGTGTGCAAAGATGCCACAGTGCACAAACCCACCCACATCTGCAT	602
Db	756	TGTGATGTGGAGAGTGTGCAAAGATGCCACAGTGCACAAACCCACCCACATCTGCAT	815
QY	603	CAATCTGAAGCGGGTACACCTGCTCTGCACGACGAGATTTGGCTTCTGGAAGGCCA	662
Db	816	CAATCTGAAGCGGGTACACCTGCTCTGCACGACGAGATTTGGCTTCTGGAAGGCCA	875
QY	663	GTGCTTAGACATGTGATGTGCGTATGTGTTACGTGCAGGAGCTGTGGCAATGTTCC	722
Db	876	GTGCTTAGACATGTGATGTGCGTATGTGTTACGTGCAGGAGCTGTGGCAATGTTCC	935
QY	723	TGGATCTATCTTGTGTACATGCAACCCGTGTTTACCCCTCATATGAGATGGAAGGCTCTTG	782
Db	936	TGGATCTATCTTGTGTACATGCAACCCGTGTTTACCCCTCATATGAGATGGAAGGCTCTTG	995
QY	783	CCAAGATGTGAACGAGTGTGCCACCGAGAAACCCCTGCTGCAAACTCGCTCAACACTTA	842
Db	996	CCAAGATGTGAACGAGTGTGCCACCGAGAAACCCCTGCTGCAAACTCGCTCAACACTTA	1055
QY	843	CGGCTCTTTCATCTGCGCTGTGCACCCAGATATGAATGAGGAAGATGGCGTTCATTTG	902
Db	1056	CGGCTCTTTCATCTGCGCTGTGCACCCAGATATGAATGAGGAAGATGGCGTTCATTTG	1113
QY	903	CAGTGATGTGACGAGTGCAGCTTCTCTGAGTTCCTGTGCCAACATGAGTGTGAAACCA	962
Db	1116	CAGTGATGTGACGAGTGCAGCTTCTCTGAGTTCCTGTGCCAACATGAGTGTGAAACCA	1175
QY	963	GCCGGGACAATCTTCTGCTCTGCGCCCTCGAGGCTACATCTCGGTGGATGGAACCCGAG	1022
Db	1176	GCCGGGACAATCTTCTGCTCTGCGCCCTCGAGGCTACATCTCGGTGGATGGAACCCGAG	1235
QY	1023	CTGCGCAAGACATCAAGATGTGAGCACAGGAAACACAGTGCACACCTGCGAGACGCTG	1087
Db	1236	CTGCGCAAGACATCAAGATGTGAGCACAGGAAACACAGTGCACACCTGCGAGACGCTG	1299
QY	1083	CTACAATTTCAAGAGGGGCTTCAAAATGATGAGACCCCATCCGCTGTGAGGAGCCTTATCT	1142
Db	1296	CTACAATTTCAAGAGGGGCTTCAAAATGATGAGACCCCATCCGCTGTGAGGAGCCTTATCT	1355
QY	1143	GAGGATAGATTAACCGCTGTATGTGCTGCTGAGAAACCTGCGTGCACAGACACAGCC	1202
Db	1356	GAGGATAGATTAACCGCTGTATGTGCTGCTGAGAAACCTGCGTGCACAGACACAGCC	1415
QY	1203	CTTTACCATCTTGTACCGGGACATGACGTGTGTGTCAGAGACGCTCGTTCGCCGTGACAT	1262

Matches 2325; Conservative 0; Mismatches 1; Indels 0; Gaps 0;				
QY	3	CCCGGGCTCTCCCGTCTCTCCACGACTCGTCGGCCCTCTGTGAATAAACACCC	62	
Db	216	CCCGGGCTCTCCCGTCTCTCCACGACTCGTCGGCCCTCTGTGAATAAACACCC	275	
QY	63	GGGAGCCCCGAGGGCCAGAGAGAGCGACGTGCCGAGCTCTCCGGGGTCCCGCCG	122	
Db	276	GGGAGCCCCGAGGGCCAGAGAGAGCGACGTGCCGAGCTCTCTCCGGGGTCCCGCCG	335	
QY	123	CGAGCTTCTCTCCGCTTGGCATCTCTCTCGCGCTCTTGGACATGCCAGATAA	182	
Db	336	CGAGCTTCTCTCCGCTTGGCATCTCTCTCGCGCTCTTGGACATGCCAGATAA	395	
QY	183	AAGGATCTACCTGTACATTCCTGTCTCTTCCAGCCCTGGGAATGCACAGC	242	
Db	396	AAGGATCTACCTGTACATTCCTGTCTCTTCCAGCCCTGGGAATGCACAGC	455	
QY	243	ACAGTCACGAAATGCTTTGACCTGATTCGCGACATCAGACAGTGTATTTAGATGA	302	
Db	456	ACAGTCACGAAATGCTTTGACCTGATTCGCGACATCAGACAGTGTATTTAGATGA	515	
QY	303	ATGCCGAAACATCCCGAGGCTGGCGAGAGACATGTTGTATACCAAAATGGCG	362	
Db	516	ATGCCGAAACATCCCGAGGCTGGCGAGAGACATGTTGTATACCAAAATGGCG	575	
QY	363	GTATTTATGCAATCCCGGAGCAAAACCTGTGTATCGAGGGCCCTACTCGAACCCCTACTC	422	
Db	576	GTATTTATGCAATCCCGGAGCAAAACCTGTGTATCGAGGGCCCTACTCGAACCCCTACTC	635	
QY	423	GACCCCTACTCAGTCCGTAACCCAGACGTGCCCAACCACTCTCAGCTCCAACTATCC	482	
Db	636	GACCCCTACTCAGTCCGTAACCCAGACGTGCCCAACCACTCTCAGCTCCAACTATCC	695	
QY	483	CACGATCTCAGGCTCTTATATGCCCTTGTGTATGATACAGATGATGAAGAACCATG	542	
Db	696	CACGATCTCAGGCTCTTATATGCCCTTGTGTATGATACAGATGATGAAGAACCATG	755	
QY	543	TGTGATGTGAGAGAGTGTCAACAGATTCACACAGTGCACACCCACAGATCTGCAT	602	
Db	756	TGTGATGTGAGAGAGTGTCAACAGATTCACACAGTGCACACCCACAGATCTGCAT	815	
QY	603	CAATCTGAAGCGGGGTACCTGCTCTCTGCACCGAGGATATTTGGCTTCTGGAAGCCA	662	
Db	816	CAATCTGAAGCGGGGTACCTGCTCTCTGCACCGAGGATATTTGGCTTCTGGAAGCCA	875	
QY	663	GTGCTTAGACATTTGATGAATGTGCTATGTTACTGTGCACAGAGCTCTGTGGAAATGCC	722	
Db	876	GTGCTTAGACATTTGATGAATGTGCTATGTTACTGTGCACAGAGCTCTGTGGAAATGCC	935	
QY	723	TGATCTCTATTTCTTTACATGCAACCTGTTTACCTCAATGAGATGGAAGTCTTG	782	
Db	936	TGATCTCTATTTCTTTACATGCAACCTGTTTACCTCAATGAGATGGAAGTCTTG	995	
QY	783	CCAAGATGTGACAGAGTGTGCCACCGAAGAACCCCTGGCTGCAAACTGCTGACACCTA	842	
Db	996	CCAAGATGTGACAGAGTGTGCCACCGAAGAACCCCTGGCTGCAAACTGCTGACACCTA	1055	
QY	843	CGGCTTTTCACTGTGCGCTGTGACCGAGATATGAATGAGGAAGATGGGCTCATTTG	902	
Db	1056	CGGCTTTTCACTGTGCGCTGTGACCGAGATATGAATGAGGAAGATGGGCTCATTTG	1115	
QY	903	CAGTGATGTGACAGAGTGTCTCTGTGATTCCTCTGCGCAACATGATGATGTAACCA	962	
Db	1116	CAGTGATGTGACAGAGTGTCTCTGTGATTCCTCTGCGCAACATGATGATGTAACCA	1175	
QY	963	GGCCGGCACATCTTCTGCTCTGCGCTCCAGGCTACATCTGCTGTGATGACCAAGG	1022	
Db	1176	GGCCGGCACATCTTCTGCTCTGCGCTCCAGGCTACATCTGCTGTGATGACCAAGG	1235	
QY	1023	CTGCCAAGACATGACGAATGTGAGCAGAGAACCAACGTCGAACCTGACGACGACGTCG	1082	
Db	1236	CTGCCAAGACATGACGAATGTGAGCAGAGAACCAACGTCGAACCTGACGACGACGTCG	1295	

QY	1083	CTACAATTTTCAAGGGGGCTTCAATGATGACGACCCCATCCGCTGTGAGAGCTTATCT	1142	
Db	1296	CTACAATTTTCAAGGGGGCTTCAATGATGACGACCCCATCCGCTGTGAGAGCTTATCT	1355	
QY	1143	GAGGATCAGTATTAACCGCTGTATGTCTCTGTGAGAACCTGGCTGCAGACGACCC	1202	
Db	1356	GAGGATCAGTATTAACCGCTGTATGTCTCTGTGAGAACCTGGCTGCAGACGACCC	1415	
QY	1203	CTTTACATCTGTACCGGAGCATGAGACGTGTGACAGACGCTCCGCTGACAT	1262	
Db	1416	CTTTACATCTGTGTACCGGAGCATGAGACGTGTGACAGACGCTCCGCTGACAT	1475	
QY	1263	CTTCAATGCAAGCCACGACCCGCTACCTCTGGGGCTATTACATTTTCCAGATCAATC	1322	
Db	1476	CTTCAATGCAAGCCACGACCCGCTACCTCTGGGGCTATTACATTTTCCAGATCAATC	1535	
QY	1323	TGGGAATGAGGCGCAGAAATTTACATGCGGCAAAAGGGCCCATCAGTCCACCTGGT	1382	
Db	1536	TGGGAATGAGGCGCAGAAATTTACATGCGGCAAAAGGGCCCATCAGTCCACCTGGT	1595	
QY	1383	GATGACAGGCCCCCTCAAGGGGCCCGGGAATCCAGCTGAGCTGGAATGATCACTGT	1442	
Db	1596	GATGACAGGCCCCCTCAAGGGGCCCGGGAATCCAGCTGAGCTGGAATGATCACTGT	1655	
QY	1443	CAACACTGTATCAACTTCAGAGGACGCTCCGTATCCGACATGCGGATATATGTGCGA	1502	
Db	1656	CAACACTGTATCAACTTCAGAGGACGCTCCGTATCCGACATGCGGATATATGTGCGA	1715	
QY	1503	GTACCCATTTGAGCTTCGGGCTGAGCTCCGACGCTGCTCTCATTTGGCACCAAGGA	1562	
Db	1716	GTACCCATTTGAGCTTCGGGCTGAGCTCCGACGCTGCTCTCATTTGGCACCAAGGA	1775	
QY	1563	CAGGAGAAGAGGAATTAACAGAGAAATAGAGCGACACAGAGTTAGCATTTCCG	1622	
Db	1776	CAGGAGAAGAGGAATTAACAGAGAAATAGAGCGACACAGAGTTAGCATTTCCG	1835	
QY	1623	CTGAAGCTTTCGCCGAGAGAGTCAACCCGACTTCTGTGACTGCACCTGACTATTGCGA	1682	
Db	1836	CTGAAGCTTTCGCCGAGAGAGTCAACCCGACTTCTGTGACTGCACCTGACTATTGCGA	1895	
QY	1683	CTGTGACCTGACGAGACTTCCACCCAGTTCCTATGATATGATGATCAAAAGTAT	1742	
Db	1896	CTGTGACCTGACGAGACTTCCACCCAGTTCCTATGATATGATGATCAAAAGTAT	1955	
QY	1743	ATCATCTCCCTCGATGAGAAATGTTGGGAATTTCAAGGCTTCAGTTATTTCCA	1802	
Db	1956	ATCATCTCCCTCGATGAGAAATGTTGGGAATTTCAAGGCTTCAGTTATTTCCA	2015	
QY	1803	CTATTTTCAAAAGAAATAGATTAGTTTGGCGGGCTGAGTCTATGTTCAAAAGACTGTG	1862	
Db	2016	CTATTTTCAAAAGAAATAGATTAGTTTGGCGGGCTGAGTCTATGTTCAAAAGACTGTG	2075	
QY	1863	AACAGCTGCTGTACCTTCTTACCTCTCTCCACTCTCTCTCTCTCTCTCTCTCTCT	1922	
Db	2076	AACAGCTGCTGTACCTTCTTACCTCTCTCTCCACTCTCTCTCTCTCTCTCTCTCTCT	2135	
QY	1923	CAAAAGACCGGGAGCTGGCGGGGAACCTGGGAGATGCTAGTTTCTTTTCTGACACA	1982	
Db	2136	CAAAAGACCGGGAGCTGGCGGGGAACCTGGGAGATGCTAGTTTCTTTTCTGACACA	2195	
QY	1983	GAGAAGCTATGTAACAAACACAGCAGATGCAAGGGTTTTAGAGAAATGTGTTCAA	2042	
Db	2196	GAGAAGCTATGTAACAAACACAGCAGATGCAAGGGTTTTAGAGAAATGTGTTCAA	2255	
QY	2043	AACCATGCTGTGATTTTCAACATTAAGAAAGTTTCAAGTTTCTTAAATTTGATTAAC	2102	
Db	2256	AACCATGCTGTGATTTTCAACATTAAGAAAGTTTCAAGTTTCTTAAATTTGATTAAC	2315	
QY	2103	GGTTAATCTGTCTGTGATTTTCAATTTTAAATTTTAAATTTTAAATTTTAAATTTT	2162	
Db	2316	GGTTAATCTGTCTGTGATTTTCAATTTTAAATTTTAAATTTTAAATTTTAAATTTT	2375	

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 10:57:33 ; Search time 107.424 Seconds

(without alignments)
6646.008 Million cell updates/sec

Title: US-09-674-379A-12

Perfect score: 2328
Sequence: 1 gaccgcgcgcctccctccgty.....tgattaatattgaagaga 2328

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2.6/ptodata/1/lna/5A.COMB.seq:*\n2: /cgn2.6/ptodata/1/lna/5B.COMB.seq:*\n3: /cgn2.6/ptodata/1/lna/6A.COMB.seq:*\n4: /cgn2.6/ptodata/1/lna/6B.COMB.seq:*\n5: /cgn2.6/ptodata/1/lna/PCrT5.COMB.seq:*\n6: /cgn2.6/ptodata/1/lna/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2324.4	99.8	2550	2	US-08-884-072-2
2	2324.4	99.8	2550	4	US-09-212-168-2
3	896	38.5	1260	3	US-09-188-930-67
4	386	16.6	1531	2	US-08-833-963C-1
5	386	16.6	2018	3	US-08-980-514-2
6	331.4	14.2	1018	3	US-09-188-930-259
7	241.4	10.4	2312	4	US-09-248-757-1
8	106.8	4.6	3373	2	US-08-897-443-2
9	89.4	3.8	401	4	US-09-643-597-278
10	80.2	3.4	220	4	US-09-404-879A-340
11	58.2	2.5	2461	1	US-08-282-141-1
12	53.2	2.3	3759	3	US-08-479-722B-3
13	52.4	2.3	5499	3	US-08-479-722B-1
14	52.4	2.3	5502	5	PCT-US95-02251-17
15	51.6	2.2	3753	5	PCT-US95-02251-2
16	51.6	2.2	4314	1	US-08-199-780-2
17	51.6	2.2	4314	1	US-08-316-650-2
18	51.6	2.2	5089	6	5177197-31
19	50.4	2.2	6412	4	US-08-652-877-17
20	50.4	2.2	6412	4	US-08-476-515A-17
21	50.4	2.2	14044	4	US-08-652-877-85
22	50.4	2.2	14044	4	US-08-652-877-89
23	50.4	2.2	14080	4	US-08-652-877-87
24	50.4	2.2	14083	4	US-08-476-515A-83
25	50.4	2.2	14086	4	US-08-652-877-83
26	45.8	2.0	7218	1	US-08-232-463-14
27	45.6	2.0	3460	2	US-08-751-305-1

28	44.8	1.9	2181	6	5208144-36	Patent No. 5208144
29	43.4	1.9	3546	4	US-08-872-757-3	Sequence 3, Appli
30	41.4	1.8	2492	4	US-09-381-779-1	Sequence 1, Appli
31	39.8	1.7	3156	4	US-09-284-819-8	Sequence 8, Appli
32	39.6	1.7	1639	2	US-08-468-819-77	Sequence 77, Appli
33	38	1.6	6048	4	US-09-634-920-3	Sequence 3, Appli
34	37.8	1.6	1611	4	US-09-249-697A-2	Sequence 2, Appli
35	37.8	1.6	1611	4	US-09-363-316B-2	Sequence 2, Appli
36	37.8	1.6	2365	4	US-09-249-697A-5	Sequence 5, Appli
37	37.8	1.6	2365	4	US-09-249-697A-18	Sequence 18, Appli
38	37.8	1.6	2365	4	US-09-363-316B-5	Sequence 23, Appli
39	37.8	1.6	2365	4	US-09-363-316B-23	Sequence 23, Appli
40	37.2	1.6	2397	6	5258288-2	Patent No. 5258288
41	37	1.6	6344	4	US-08-843-417-1	Sequence 1, Appli
42	36.6	1.6	5099	4	US-09-610-040-5	Sequence 5, Appli
43	36.6	1.6	72328	3	US-09-009-913-1	Sequence 1, Appli
44	35.4	1.5	825	1	US-08-312-870-6	Sequence 6, Appli
45	35.4	1.5	1338	1	US-08-307-444A-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-08-884-072-2
Sequence 2, Application US/08884072
Patent No. 5872234
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Guebler, Karl J.
TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/884,072
FILING DATE: Herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0333 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2550 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: CORNOT01
CLONE: 45517
US-08-884-072-2
Query Match 99.8%; Score 2324.4; DB 2; Length 2550;
Best Local Similarity 100.0%; Pred. No. 0;

OY	664	GGCTTAGACATTGAGTANGTGCGATGGTATCTGACGACGACGCTGTGGCAATGTCTCT	723
Dp	766	TGCTTAGATATTGAGTAATGTGCTATGGTATCTGACGACGCTGTGGCAATGTCTCA	825
OY	724	GGATTCCTATTTCTGTACATGCAACCCCTGGTTTTACCCCAATAGAGATGGAAGCTTTGC	783
Dp	826	GGATTCCTATTTCTGTACATGCAACCCCTGGTTTTACCCCAATAGAGATGGAAGCTTTGC	885
OY	784	CAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCGTGAACACCTGCGTCAACACCTAC	843
Dp	886	CAAGATGTGAACGAGTGTGCCAAACTGAGATCCCTGTGTACAGACCTGTGTCAACACCTAT	945
OY	844	GGCTTTTCATCTGGCGGTGTGACCCAGAAATATGACATTTGAGGAGATGGCGTTATATGC	903
Dp	946	GGCTTTTCATCTGGCGGTGTGACCCAGAAATATGACATTTGAGGAGATGGCGTTATATGC	1005
OY	904	AGTGAATGAGAGAGTGCAGCTTCTCTGAGTTCTCTGCTCCCAACATGAGTGTGTGAACAG	963
b	1006	AGTGAATGAGAGAGTGCAGCTTCTCTGAGTTCTCTCTCTCAACAGAGTGTGTGAACAG	1066
OY	964	CCCGGCACATACCTTCTGCTCCTGCCCTCCAGGCTACATCTGTGTGATGACAAACGAGC	1022
Dp	1066	CCGGGCTATATCTTCTGCTCGTGCCCTCCAGGCTACATCTGTGTGATGATACGGAAGC	1122
OY	1024	TGCCAAGACATCAACGAATGTGACACAGGAACACACAGTGCACCTGCAGACAGCTGC	108
Dp	1126	TGCCGAGTATCAAAAGAAATGTGAGACCCGAAACCAACAGTGTACTCTCAGTCAAGCTTGC	1185
OY	1084	TACATATTACAAGGGGGCTTCAAAATGACATGCACCCCATCCGCTGTGAGAGAGCTTATG	1144
Dp	1186	TACATATTACAAGGGGGCTTCAAAATGTATTGTCCATCAGTGTGAGAGAGCTTATG	1245
OY	1144	AGGATCACTGTAACCCGCTGTATGTCTCTGCTGAGAACCCCTGGCTGTGAGAGACCGCC	1203
Dp	1246	CTGATTTGGTGAAAACCGCTGTATGTCTCTGCTGTAGACACACAGTGTGAGAGACCGCA	1305
OY	1204	TTTACCATCTTGTACCGGGACATGGACGTGGTGTACAGACGCTCCGTTCCGCTGCATAC	126
Dp	1306	TTTACCATCTTGTATCGGGACATGGATGTGGTGTGTCAGGACGCTCCGTTCCGCTGCATAC	1365
OY	1264	TTTCCAAATGCAAGCAGCACCGCGCTGAGGCGCTTATTCATTTTCAGATCAAAACT	1322
Dp	1366	TTTCCAGATGCAAGCAACACCGGATACCTGTGTGCTTATTCATTTTCAGATCAAAACT	1422
OY	1324	GGGAATGAGGCGAGAAATTTTACATGCGGAAACGGGCCCCCATCACTGACCAACCTGGTG	1383
Dp	1426	GGCAACGAGGGGTGAGAACTTGTATATGTGGCAAAACAGGGCGCTATGATGCCACCTGGTG	1485
OY	1384	ATGACACGCCCCCATCAAAAGGCCCCGGGAAATCCAGCTGACCTTGTGAATATGATCACTGC	1443
Dp	1486	ATGACACGCCCCCATCAAAAGGCCCCGTGAGCACTCAGCTGTGGAATATGATCACTGC	1543
OY	1444	AACACTGTCAATCACTTCAAGGAGGACGCTCGATACCGACGCGGATTTATGTGTGCGAG	1503
Dp	1546	AACACTGTCAATCACTTCAAGGAGGACGCTCGATACCGACGCGGATTTATGTGTGCGAG	1605
OY	1504	TACCATTTCTAGAGCTCGGGCTGTGAGGCTCCGACGCTCCTCTCATTTGACCAAGGGAC	1563
Dp	1606	TATCCGTTCTAGAGCTCTGGGCTTAAGGCTCTGTACACTGCTTTTCAACGACCGAGGGAC	1666
OY	1564	AGGAGAGAGAGAAATTAACAGAGAGAAATGAGAGGACACAGACGTTTAGGACTTCTGC	1622
Dp	1666	GGGAGAGAGAGAAAC--CAGCAAGAAATGAGAGGAGAGACATTCACACTTTCCTGC	1722
OY	1624	TGAACGTTTCCCCGAAAGATGACGCCGACCTTCTGTACTCTACCTGTACTATTGTGACAG	1688
Dp	1724	TGAATATCTCTCTGGGGGATCAGGCTAG-CATCTTGAACCCATATCTGTACTATTGACAG-	1781
OY	1684	CTGTACACCTGACAGACTTGCACCCCCAGTTTCCATATATACAGTTATCAAAAAGTATTA	1743
Dp	1782	TGTGTACTCTTAAGAGCACCTCGCCCTGTAGTTCTATGATATCAGATTATCCAAAAGTGTTC	1841

QY	1744	TCATTCGCCCCCTGATGAGAAATGTTGGTCGTAATTTCCAGGCCCTTCAGTTATTTCCAC	1803
Db	1842	ATCTTT-ACCCTCTGATATGAGGTTGCCAGTACCTCTCAAAAGCCCTTCATTATTTCCAT	1900
QY	1804	TATTTT----CAAGAATAATAGATTAGTTGGCGGGGGTCCGAG----TCGTATCTCAAG	1856
Db	1901	CGTTTATATAAAAAAATAATAGATTAGATTGGCTGGGGTATAGATCTCCAGAGTTCCAA	1960
QY	1857	ACTGTGAACAGCTTGCCTGTCACTTCTTCACCTTCCACATCTCCTCTCACTGTGTTACT	1916
Db	1961	AGAGCTGAGTGGCTTGCCTCTCACACCTTCCTCTCCTTC-CTCCACATCTTGGCTGATTGCT	2019
QY	1917	GCTTTGCAAAACCCCG--GAGCTGGGGGGGAACCCGGGAGTGGAGTTGCTTTTG	1974
Db	2020	GCTTTGCAAAAGTCTCTCATGGGGCTCGTGGGAAAGCTGGGAAATGATGATTTGCTTTTG	2079
QY	1975	CGTACACAGAAAGGCTATGTAAACAAACACACAGAGATGAAGGGTTTTTATAGATG	2034
Db	2080	CATGTTCTGAGAAAGGCTATGGGAAACACACACAGAGATCGAA-GGTTTTTATAGAGTC	2138
QY	2035	TGTTTCAAAACCATGCCTGTATTTTCAACCATAAAGAAAGTTTCACTGTCTTAAATT	2094
Db	2139	TATTTTAAAAACATCATCTCGATTTTTCAGCATAAAGAAATTTAGTTGTCCTTAAATT	2198
QY	2095	TGTTATACGGTTTAAATCTGCTCTGTTCATTTTAGATATTTTAAAAAATATGTCGTAGA	2154
Db	2199	TGTTATGAGTGTTTAAACCTTTCTTAAATTCATTTTAGGCTCTCTTAA-----GGGTAGA	2252
QY	2155	ATTCCCTCGAAAGGCTTCACAGACATGCTATGTTGTCTTCCCAAAACCAAGCTCCTC	2214
Db	2253	ATTCTTTCCAAAGGCC-TCAGATACATGTTATGTTTCAGTCTTTCCAAACCTCATCTTTCC	2311
QY	2215	TCCATTTTTCAGCCAGTGTCTTTCTTTCAGGAGCCCTTATCTTCTTTCTTGAATTTT	2274
Db	2312	TGCATCTTAGCCCAAGTTTTTAC-----GAAAGACCCCTTATCATGCTTTNTTTPAAGAGTTTT	2368
QY	2275	ACCCATATGATTTGGATTCGAGAGTCTCCCAAATGATTAATTTTGGAAAGA	2328
Db	2369	ACCCAACTGCTTTGGAAAGACAGAGTATCCACAGATGATTAATTAATTAAGAGAAA	2422

Search completed: July 3, 2003, 11:59:30
Job time : 522.747 secs

```

QY 1504 TACCATTCTGAGCCTCGGGCTGGAGCCTCCGACGCTCTCTCATTTGGACCAAGGAC 1563
  || || || || || || || || || || || || || || || || || || || || || ||
Db 1606 TATCCGTTCTGAGACCTCTGGCTAGGCTCTGACACTCTCTTCCACGACCAAGGAC 1665
QY 1564 AGAGAGAGAGAGAAATACAGAGAGATGAGAGACACAGACTTATAGGACTTCTCC 1623
  || || || || || || || || || || || || || || || || || || || || || ||
Db 1666 GGGAGGAGAAAGAAAC--CAGCAAGATGAGAGAGACAGACATTCACCTTCTCTCC 1723
QY 1624 TGAAGCTTCCCGAAGAGTACGAGCCGAGCTCTGATCTCACTCTACTTCTTCCAGAC 1683
  || || || || || || || || || || || || || || || || || || || || || ||
Db 1724 TGAATATCTCTGGGGGATCAGCTAG--CATCTTGCCCATCTCTGATCTTCTCAGT- 1781
QY 1684 CTGTCAACCTTGACAGACTTCCACCCAGTCTCTATGATACAGTTATCAAAAGATATA 1743
  || || || || || || || || || || || || || || || || || || || || || ||
Db 1782 TGCTCACTCTGAGAGACACCTGCTCCAGTCTCTATGATACAGTTATCAAAAGATATA 1841
QY 1744 TCATGCTCCCTGATGAGAGATGTTGGTGAATTTTCAGAGCTTCAAGTTATTTCCAC 1803
  || || || || || || || || || || || || || || || || || || || || || ||
Db 1842 ATCTT-AGCCCTGATGAGGTTGCCAGTCACTCTCAAAAGCTTCCATTTATTTCCAT 1900
QY 1804 TATTTT---CAAGAAATAGATAGTTTGGGGGGGCTGAG--TCTATGTCGAAG 1856
  || || || || || || || || || || || || || || || || || || || || || ||
Db 1901 CGTTTATATAAAAGAAATAGATAGTTTGGGGGATGAGCTCTGAAAGTTCAAA 1960
QY 1857 ACTGTGAACGCTTGTCTCACTTCTTCACTCTTCTCTCTCACTGTGTACT 1916
  || || || || || || || || || || || || || || || || || || || || || ||
Db 1961 AGACTGAGTGGCTGTCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2019
QY 1917 GCTTTCGAAGACCCGG--GAGCTGGGGGGGAGACCTTGGGAGTACTAGTTGCTTTTG 1974
  || || || || || || || || || || || || || || || || || || || || || ||
Db 2020 GCTTTCGAAGAGTCTCATGGGCTGAGGAAATGCTGGAAATAGCTAGTTGCTTCTTG 2079
QY 1975 CGTACACAGAGAGGCTATGTAAACAAACACAGAGAGTGAAGGTTTATAGAAATG 2034
  || || || || || || || || || || || || || || || || || || || || || ||
Db 2080 CATGTTCTGAGAGGCTATGGAACACACACAGAGATGGA--GTTTTATAGAGTC 2138
QY 2035 TGTTCACAAACATGCTGTATTTTCAACATTAAGAGTTCAGTTGCTTAAAT 2094
  || || || || || || || || || || || || || || || || || || || || || ||
Db 2139 TATTTTAAATACATCTGTATTTTCAAGCAATTAAGAAATTTAGTTGCTTAAAT 2198
QY 2095 TGTATACGCTTAAATCTGCTGTCTCATTTTGAATTTTAAATAATGCTGTA 2154
  || || || || || || || || || || || || || || || || || || || || || ||
Db 2199 TGTATGAGTGTAAACCTTCTTATTCATTTTGAAGCTTCTTAA-----GTGGTAA 2252
QY 2155 ATTCCTTGAAGGCTTTCAGACACATGCTATGTTCTTCCAAACCCAGTCTCTC 2214
  || || || || || || || || || || || || || || || || || || || || || ||
Db 2253 ATTCCTTGAAGGCTTTCAGACACATGCTATGTTCTTCCAAACCCAGTCTCTC 2311
QY 2215 TCCATTTTACCCAGTGTCTTCTTGAAGACCCCTTAAATCTTCTTGAATTTT 2274
  || || || || || || || || || || || || || || || || || || || || || ||
Db 2312 TGCATCTTACCCAGTGTCTTGAAGACCCCTTAAATCTTCTTGAATTTT 2368
QY 2275 ACCCAATGATGAGATGACAGAGTCTCCAAACATGATTAATATTTGAAGGA 2328
  || || || || || || || || || || || || || || || || || || || || || ||
Db 2369 ACCCACTGCTGGAAGACAGAGTATCCAGACTGATTAATATTTGAAGGA 2422

```

RESULT 15

AAZ40031 standard; DNA: 2429 BP.

AAZ40031;

15-FEB-2000 (first entry)

Full length mouse A55 protein coding sequence.

A55 protein; mouse; smooth muscle proliferation; tissue generation;
 vascular smooth muscle cell; arteriosclerosis; tissue repair; myoma;
 vascular endothelial thickening; haematopoietic cell-regulation; cytokine;
 percutaneous transluminal coronary angioplasty; blood coagulation; PTCA;
 actin; inhibitor; chemotaxis; thrombosis; cadherin; therapy;
 tumour metastasis inhibitor; ss.

```

XX Mus musculus.
OS
PN MO9955864-A1.
XX
PD 04-NOV-1999.
XX
PF 28-APR-1999: 99WC-JP02284.
XX
PR 28-APR-1998: 98JP-0119731.
XX
PA (ONOX) ONO PHARM CO LTD.
PI Honjo T, Tashiro K, Nakamura T;
XX
DR MPI; 2000-038647/03.
DR P-PSDB; AAY54991.
XX
PT Novel human polypeptides for treatment of, e.g. arteriosclerosis and
  myoma -
XX
PS Example 5: Page 64-68; 87pp; Japanese.
XX
CC This sequence encodes the mouse A55 protein. The invention relates to the
CC human A55 protein. The protein can be used for the treatment of diseases
CC due to abnormal proliferation of smooth muscle. The polypeptides can be
CC used according to their inhibition of the proliferation of vascular smooth
CC muscle cells, particularly in treating arteriosclerosis or re-narrowing
CC by vascular endothelial thickening after percutaneous transluminal
CC coronary angioplasty (PTCA), or myoma, haematopoietic cell-regulatory
CC activity, cytokine activity, tissue generation/repair activity, blood
CC coagulation/thrombotic activity, receptor/ligand activity,
CC cadherin/tumour metastasis inhibiting activity; tumour inhibition, and as
CC a nutrient.
XX
SQ Sequence 2429 BP; 610 A; 611 C; 562 G; 645 T; 1 other;
XX
Query Match 62.3%; Score 1449.8; DB 21; Length 2429;
  Best Local Similarity 83.3%; Pred. No. 0;
  Matches 1795; Conservative 0; Mismatches 333; Indels 26; Gaps 12;
QY 184 AGATACCTCACTGTTACCATTTGCTCTCTCTCTTCCAAAGCTTGGGAATCCACAGCA 243
  || || || || || || || || || || || || || || || || || || || || || ||
Db 286 ATGATACCTCACTGTTACCATTTGCTCTCTCTCTTCCAAAGCTTGGGAATCCACAGCA 345
QY 244 CAGTGCACGAATGCTTTGACCTGATGCGCAGTACAGACATGTTTATGATGAA 303
  || || || || || || || || || || || || || || || || || || || || || ||
Db 346 CAGTGCACGAATGCTTTGACCTGATGCGCAGTACAGACATGTTTATGATGAA 405
  || || || || || || || || || || || || || || || || || || || || || ||
QY 304 TGGCGAACCATCCCGAGGCTTGGCGAGAGACATGATGTTTAAACCAATGGCGGG 363
  || || || || || || || || || || || || || || || || || || || || || ||
Db 406 TGGCGAACCATCCCGAGGCTTGGCGAGAGACATGATGTTTAAACCAATGGCGGG 465
  || || || || || || || || || || || || || || || || || || || || || ||
QY 364 TATTTATGATTTCCCGAGCAACACCTGTGTATGAGGAGGCTTCTGAAACCTCACTG 423
  || || || || || || || || || || || || || || || || || || || || || ||
Db 466 TATTTATGATTTCCCGAGCAACACCTGTGTATGAGGAGGCTTCTGAAACCTCACTG 525
QY 424 ACCCCCTACATCAGTTCGCTACCCAGAGCTGCTCCACACATCTCAGTCCAAATGCC 483
  || || || || || || || || || || || || || || || || || || || || || ||
Db 526 ACATCTCACTCAGTTCGCTACCCAGAGCTGCTCCACACATCTCAGTCCAAATGCC 585
  || || || || || || || || || || || || || || || || || || || || || ||
QY 484 ACATCTCACTCAGTTCGCTACCCAGAGCTGCTCCACACATCTCAGTCCAAATGCC 543
  || || || || || || || || || || || || || || || || || || || || || ||
Db 586 ACATTTCAAGGCTCTTGTCTGCGCGCTTTGGGTATCAGATGATGAAAGCAACCACT 645
  || || || || || || || || || || || || || || || || || || || || || ||
QY 544 GTGATGTGACAGAGTGTCAACAGATTTCCACAGTCAACCCACCCAGATCTGATC 603
  || || || || || || || || || || || || || || || || || || || || || ||
Db 646 GTGATGTGACAGAGTGTCAACAGATTTCCACAGTCAACCCACCCAGATCTGATC 705
  || || || || || || || || || || || || || || || || || || || || || ||
QY 604 AATACCTAAGCGGGTACACTGCTCTGACACCAAGCATATGCTTCTGAAAGCCAG 663
  || || || || || || || || || || || || || || || || || || || || || ||
Db 706 AACACTGAAGAGGTACACTGCTCTGACACCAAGCATATGCTTCTGAAAGCCAG 765

```


223 ATGCACAGCAGATGCGACGATGCTTGTGACGTGATCCGACGAGCAGTGTAG 292
139 ATGCACAGCAGCAGTGCACAAAGGCTTTGACCTGGACCGCAGTGCAGACAGTCTAG 198
293 ATATTGATGAAATCCGAAACCATCCCGAGCGCTCCGAGAGACATGATGTGTAAAC 352
199 ATATTGATGAAATCCGAGACCATCCCTGAGGCTTGTGAGGACATGATGTGTCAAC 258
353 AAAATGCGGGGTATTTATGATTCGCCGAGCAAAACCTGTGTATGAGAGGCGCTACTGA 412
259 AGATGCGGGGTATTTGTGTGATCCCTGAAACCAACAGTGTATGAGAGGCGCTTACTAA 318
413 ACCCTGCTGACAGCCCTTACTAGGTCCGTACCCAGAGAGTGCCTCCACCTCTGAGTC 472
319 ATCCCTACTCTATCTTACTAGGCCCATTCACAGACGCGGCCACACAGTACAGCTT 378
473 CAATGATCCACAGATCTCCAGGCGCTGTATATGCCGCTTGTGATACAGATGATGAAA 532
379 CCACCTACCCACAGATTCAGGCGCTGTGTGCTGCCCTTGGGTATGATGATGAAAG 438
533 GCAACCAATGTGTGATGAGTGGAGTGGCAAGATTCGCCACAGTGCACACCCACCC 592
439 GCAACCAAGTGTGTGATGAGTGGAGTGGCAAGATTCGCCACAGTGCACACCCACCC 498
593 AGATCTGATCAATGATGAGAGCGGTGACACCTGCTCCAGCAGCAGATATGGCTTC 652
499 AGATCTGATCAATGATGAGAGCGGTGACACCTGCTCCAGCAGCAGATATGGCTTC 558
653 TGGAGGCGCAGTCTTACATGATGATGATGATGATGATGATGATGATGATGATGATG 712
559 TGGAGGCGCAGTCTTACATGATGATGATGATGATGATGATGATGATGATGATGATG 618
713 CGAATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 772
619 CAATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 678
773 GAAGGTCTTGGCAAGATGTAAGCAGTGTGCACCGAGAACCCCTGCTGCAAAACCTGCG 832
679 GAAGGTCTTGGCAAGATGTAAGCAGTGTGCACCGAGAACCCCTGCTGCAAAACCTGCG 738
833 TCAACACATGAGGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 892
739 TCAACACATGAGGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 798
893 GCGTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 952
799 GCGTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 858
953 GTGTGAACCAAGCGCGGACATCTTGTGCTGCTGCCCTCAGGCTACATCTGCTGATG 1012
859 GTGTGAACCAAGCGCGGCTCTTGTGCTGCTGCCCTCAGGCTACATCTGCTGATG 918
1013 ACAACCGAAGCTTCCGACAGATATGATGATGATGATGATGATGATGATGATGATGATG 1072
919 ATTAACCGAAGCTTCCGACAGATATGATGATGATGATGATGATGATGATGATGATGATG 978
1073 ACAGAGAGTGTCAATTTTCAAGGGGGCTTCAATGATGATGATGATGATGATGATGATG 1132
979 TGAAGAGTGTCAATTTTCAAGGGGGCTTCAATGATGATGATGATGATGATGATGATG 1038
1133 AGCTTATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1192
1039 AGCTTATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1098
1193 GAGACAGGCGCTTACATCTTGTAGCGGAGCATGAGAGTGTGTGACAGAGCGCTCCGTC 1252
1099 GAGACAGGCGCTTACATCTTGTAGCGGAGCATGAGAGTGTGTGACAGAGCGCTCCGTC 1158
1253 CCGCTGACATCTTCAAAATGCAAGCAGACCGCTACCGCTGGGGCTTATTACATTTTCC 1312
1159 CTGCTGACATCTTCAAAATGCAAGCAGACCGCTACCGCTGGGGCTTATTACATTTTCC 1218
1313 AGATCAAAATCTGGGAATGAGGAGAGAAATTTTACATGCGGCAAAAGGGGCCCATCAGTG 1372

1219 AGATCAAAATCTGGGAATGAGGAGGTGACAGTCTTATATCGGCAAAAGGGGCTATCAGTG 1278
1373 CCACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1432
1279 CCACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1338
1433 TGAATCACTGTCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1492
1339 TGAATCACTGTCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1398
1493 ATGTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1552
1399 ATGTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1458
1553 CACCAAGGAGCAG 1612
1459 CACCAAGGAGCAG 1516
1613 GCATTTCTGCTGAAAGCTTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1672
1517 ACCTTCTGCTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1575
1673 CTATTTGACAGCTGCTGACCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1732
1576 CTATTTGACAGCTGCTGACCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1634
1733 AAAAAGTATATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1792
1635 CAAAAGTATATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1693
1793 TTTATTTCCACTATTTT-----CAAGAAGAAATAGATTAGTGTGCGGGGCTGAGTCT-- 1846
1694 TTTATTTCCACTATTTT-----CAAGAAGAAATAGATTAGTGTGCGGGGCTGAGTCT-- 1753
1847 -ATGTTCAAGACTGTGAACAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1905
1754 GAAGGTTCAAAAGACTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1812
1906 AGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1963
1813 GCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1872
1964 TTTGCTTTTGGCTGACAG 2023
1873 TTTGCTTTTGGCTGACAG 1931
2024 TTTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2083
1932 TTTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1991
2084 GTCTTAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2143
1992 TCTTTAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2048
2144 TATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2203
2049 ---GTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2104
2204 CCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2263
2105 TCAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2161
2264 TTTAGATTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 2323
2162 TAAAGATTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 2221
2324 AGAGA 2328
2222 AGAAA 2226

RESULT 14

Db	1099	GAGACACGCCATTTCACCATCTCTGTAATGGGACATAGATGTGTGTGCAGAGACGCTCCGTTG	1158
Oy	1253	CCGCTGACATCTTCCAAATGCAAGCCACGACCGCGTACCTCGGGGCGCTATTACATTTTCC	1312
Db	1159	CTGCTGACATCTTCCAGATGCAAGCAACACCATACCTGTGGCTATTACATTTTCC	1218
Oy	1313	AGATCAAAATCTGGGAATGAGGGGCGAGAAATTTTACATGTGGGCAAAAGGGGCCCATTCAGTG	1372
Db	1219	AGATCAAAATCTGGGAACGAGAGGGGTGAGAGTCTCTATGTCCGGCAAAACAGAGGGCTCTACGTG	1278
Oy	1373	CCACCTGTGATGACAGCGCCCATCAAGAGGGCCCCGGGAATACAGCTGTGACTTGGAAA	1432
Db	1279	CCACCTGTGTATGACAGCGCCCATCAAGAGGGCCTCGGAGACATCAGCTGGACTTTGGAGA	1338
Oy	1433	TGATCACTGTCAACACTGTCTCACTCACTTCAGAGGCAAGCTCCGTATCCGACTGGCGGATAT	1492
Db	1339	TGATCACTGTCAACACTGTCTCACTCACTTCAGAGGCAAGCTCCGTATCCGACTGGCGGATAT	1398
Y	1493	ATGTGTGGCAATACCCATCTGTGACCCCTGGGCTGGAGCTCTCCGACGCTGCTCTCATTTGG	1552
Db	1399	ATGTGTGGCAATATCCGTTCTGTGACCTCTGGCTAAAGCTCTGTGACACTGCTTTCACCG	1458
Oy	1553	CACCAAGGGACAGAGAAGAGAGAAATTAACAGAGAGATGAGAGCGACACAGACGTTAG	1612
Db	1459	CACCGAGGGACGGGAGGAGAAAGGA--ACACAGCAAGATGAGAGCGAGACAGACATTTGC	1516
Oy	1613	GCATTTCTGCTGAAACGTTTCCCGAGAGAGTCAAGCCCCGACTCTCTGACTCTACCTGTA	1672
Db	1517	ACCTTCTCCTGTAATATCTCTGGGGGCATCAGCCTAG-CATCTTGACCCATATCTGTA	1575
Oy	1673	CTATTGGACGCTGTACACCCTGACAGAGATTTGCCACCCCACTTCCATGATACAGTATATC	1732
Db	1576	CTATTGGAGA-TGGTCACTGTGAAGAGACCCCTGCCCTCACTTCTTAATGATGACAGTATTC	1634
Oy	1733	AAAAAGTATATCAATTCGCTCCCTGATGAGAGATTTGTTGGTGAATTTTCAAGCGCTTCAG	1792
Db	1635	CAAAAGTCTATCTT-AGCCCCGTATGAGGTTCGACAGTACTTTCAAAAGCTTCCA	1693
Oy	1793	TTTATTTCCACTATTTT---CAAGAAATATGATTAGGTTTGGGGGGTCTGAGTCN-	1846
Db	1694	TTTATTTCCACTATTTTATTAATAAAGAAATATGATTAGTTCGGGGGTATGAGTCTCTC	1753
Oy	1847	-ATGTCCAAGACGTGACAGACGTTTGGTGTCACTTTTCACTCTTCCACTCTCTTCTCTC	1905
Db	1754	GAAAGTTCACAAAGACTGAGTGGTGTCTCACTCTTCCCTCTCTCTC-CTCCATCTCTT	1812
Oy	1906	ACTGTGTACTAGCTTTTGGCAA--GACCGGGAGCTGGGGGGGAAACCCGGGAGTAGCTAG	1963
Db	1813	GCTGCAATTGCTGTTCGAAAAGTCTCATGTGGCTCGTGGGAAATGCTGGGAAATAGCTAG	1872
Oy	1964	TTTGTCTTTTGCCTACACAGAGAAAGGCTATGTAACAAACACAGACAGATCGAAGGGTT	2023
Db	1873	TTTGTCTTCTCATGTCTCTGAGAAGGCTATGGAACAACACACAGACAGATCGAA--GGTT	1931
Oy	2024	TTTAGAGAGATGTGTTTCAAAACCATGCTGTAATTTTCAACATTAAGAAAGTTTCACTT	2083
Db	1932	TTTATAGAGTCTATTATTAATTCACATCTGTGATTTTTCAGATTAAGAAATTTTAGTTG	1991
Oy	2084	GTCTCTAAATTTGTATAAGGTTTAATCTGTCTGTTCATTTTGAATTTTAAAAAA	2143
Db	1992	TCTTTAAATTTGTATGAGTGTTTTAACTTTTCTTATTTCAATTTTGAAGCTTCTTAAA--	2048
Oy	2144	TATGTGTAGATTTCTCTCGAAAGGCCCTTCAGACAGATCTATGTTCTGTCTCCAAAC	2203
Db	2049	---GTGTGAGAAATTCCTTCCAAAGGCC-TCAGATATCAATGATATGTTCAGTCTTCCAAAC	2104
Oy	2204	CCAGTCTCTCTCCATTTTAAAGCCAGTGTCTTCTTGAAGACCCCTTAATCTTGCTTCT	2263
Db	2105	TCATCTTCTCTGATCTTTAGCCAGTTTAACT---GAAAGACCCCTTAATCATGTGTTNT	2161
Oy	2264	TTAGATTTTAAACCAATTTGGATTGGAAATGCAAGAGTCTCCAACTGATTAATATTTTGA	2322

[illegible]

QY 2170 CTTCAGACATGCTATGTTCTGTTCCCAACCCAGTCCTCCATTTAGCCAG 2229
 |||||
 Db 1381 CTTCAGACATGCTATGTTCTGTTCCCAACCCAGTCCTCCATTTAGCCAG 1440
 QY 2230 TGTTCCTTTGAGAGACCCCTTAATCTGCTTTTGTAGAAATTTTACCCTGATGG 2289
 |||||
 Db 1441 TGTTCCTTTGAGAGACCCCTTAATCTGCTTTTGTAGAAATTTTACCCTGATGG 1500
 QY 2290 AATGAGAGGCTCCCAACGATTAATTAATTTGAAGAG 2327
 |||||
 Db 1501 AATGAGAGGCTCCCAACGATTAATTAATTTGAAGAG 1538

RESULT 12

AAZ39384

AAZ39384 standard; DNA; 2233 BP.

AAZ39384;

22-FEB-2000 (first entry)

Smooth muscle proliferation modulating protein encoding DNA.

Cell proliferation; vascular; smooth muscle cell; arteriosclerosis; PTCA;
 endothelial thickening; percutaneous transluminal coronary angioplasty;
 myoma; hematopoietic cell-regulation; cytokine; tissue generation; taxis;
 actin activity; chemotaxis; blood coagulation; thrombotic; tumor; mouse;
 metastasis; nutrient; ss.

Musculus.

MO9955863-A1.

04-NOV-1999.

28-APR-1999; 99MO-JP02283.

28-APR-1998; 98JP-0119731.

(ONOX) ONO PHARM CO LTD.

Honjo T, Tashiro K, Nakamura T;

WPI: 2000-038646/03.

P-PDB; AAY56750.

Novel mouse polypeptides for treatment, e.g. arteriosclerosis and myoma

Claim 5; Page 47-50; 70pp; Japanese.

The invention provides mouse polypeptides for treatment of diseases due
 to abnormal proliferation of smooth muscle. The polypeptides can be
 produced by standard recombinant methodology. The polypeptides can be
 used according to their inhibition of the proliferation of vascular
 smooth muscle cells, particularly in treating arteriosclerosis or re-
 narrowing by vascular endothelial thickening after percutaneous
 transluminal coronary angioplasty (PTCA), or myoma, hematopoietic cell-
 regulatory activity, cytokine activity, tissue generation/repairation
 activity, actin/inhibin activity, taxis and chemotaxis activity, blood
 coagulation/thrombotic activity, receptor/ligand activity, cadherin/
 tumour metastasis inhibiting activity; tumor inhibition, and as nutrient.
 The present sequence represents a DNA encoding the protein of the
 invention which can be used for modulating smooth muscle cell
 proliferation.

Sequence 2233 BP; 554 A; 582 C; 510 G; 586 T; 1 other;

Query Match 63.6%; Score 1480.8; DB 21; Length 2233;

Best Local Similarity 82.7%; Pred. No. 0; Mismatches 358; Indels 26; Gaps 12;

Matches 1841; Conservative 0;

QY 113 GTCCGCCCGGAGCTTCTCTTCGCTTCGATCTCCTCTCGGCGCTGTGAGATGC 172
 |||||
 Db 19 GTCCACCGGAGAGACCTCCCTTCGCGGTCGCTCTCTCCCGGCACTTGGATATGC 78
 QY 173 CAGGAATPAAAAGANACTACTGTTTACCATTTGCTCTGTCTTCCAGCCCTGGGA 232
 |||||
 Db 79 CAGGATPAAAAGANACTACTGTTTACCATTTGCTCTGTCTTCCAGCCCTGGGA 138
 QY 233 ATGCAGAGGACAGTGCAGGATGCTTTCACCTGATGCGAGTCAGAGACATGTTT 292
 |||||
 Db 139 ATGCAGAGGACAGTGCAGGATGCTTTCACCTGATGCGAGTCAGAGACATGTTT 198
 QY 293 ATATTGATGATGCGAGAACCATCCCGAGGCTGCGAGAGACATGATGTGTTTACC 352
 |||||
 Db 199 ATATTGATGATGCGAGAACCATCCCGAGGCTGCGAGAGACATGATGTGTTTACC 258
 QY 353 AAAATGGCGGATTTTATGCAATTCGCCGACAAACCTGTGTATGAGGCGCTACTCGA 412
 |||||
 Db 259 AGAATGGCGGATTTTATGCAATTCGCCGACAAACCTGTGTATGAGGCGCTACTCGA 318
 QY 413 ACCCTTACTGAGCCCTTACTGAGTCCGTACCCAGAGCTGCCCCACCACTCTCAGTTC 472
 |||||
 Db 319 ATCCCTTACTGAGCCCTTACTGAGTCCGTACCCAGAGCTGCCCCACCACTCTCAGTTC 378
 QY 473 CAACCTATCCCGACGATCTCCAGGCTTTATATGCGGCTTTGGATACGATGATGAATA 532
 |||||
 Db 379 CAACCTATCCCGACGATCTCCAGGCTTTATATGCGGCTTTGGATACGATGATGAATA 438
 QY 533 GCAACCAATGTGTGAGATGAGACGAGTGCACAGATTTCCACGATGCAACCCGACCC 592
 |||||
 Db 439 GCAACCAATGTGTGAGATGAGACGAGTGCACAGATTTCCACGATGCAACCCGACCC 498
 QY 593 AGATCTGATCAATACGTAAGCGGCTTACCTGCTCCGACCGAGGATTTGGCTTC 652
 |||||
 Db 499 AGATCTGATCAATACGTAAGCGGCTTACCTGCTCCGACCGAGGATTTGGCTTC 558
 QY 653 TGAAGGCGACGCTTATGACATTTGATGATGTGCTATGCTACTGCGACAGCTCTGTG 712
 |||||
 Db 559 TGAAGGCGACGCTTATGACATTTGATGATGTGCTATGCTACTGCGACAGCTCTGTG 618
 QY 713 CGAATGTTCCGTGATCTTATTTTACATGCAACCCGCTGTTTACCTCAATGAGATG 772
 |||||
 Db 619 CGAATGTTCCGTGATCTTATTTTACATGCAACCCGCTGTTTACCTCAATGAGATG 678
 QY 773 GAAAGTCTTCCGACGATGTAACGAGTGGCCACCGAGACCCCTGCAACCTGGG 832
 |||||
 Db 679 GAAAGTCTTCCGACGATGTAACGAGTGGCCACCGAGACCCCTGCAACCTGGG 738
 QY 833 TCAACACCTTACGCTCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGGAATG 892
 |||||
 Db 739 TCAACACCTTACGCTCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGGAATG 798
 QY 893 GCGTTTCAATGAGTATGTAACGAGTGCAGCTTCTGTGATGTTCTCTGCCAATGAGT 952
 |||||
 Db 799 GCGTTTCAATGAGTATGTAACGAGTGCAGCTTCTGTGATGTTCTCTGCCAATGAGT 858
 QY 953 GTGTGAACCGACCGCGGACATATCTGCTCTGCGCTTCCAGGCTTACCTGCTGGATG 1012
 |||||
 Db 859 GTGTGAACCGACCGCGGACATATCTGCTCTGCGCTTCCAGGCTTACCTGCTGGATG 918
 QY 1013 ACAACCGAGCTGCGAAGACATCAACGATGTGACACAGAACACACAGTGCACCTGC 1072
 |||||
 Db 919 ACAACCGAGCTGCGAAGACATCAACGATGTGACACAGAACACACAGTGCACCTGC 978
 QY 1073 AGCAGAGCTGTACAAATTTAAGGGGCTTCAATGATGATGCAACCCATCCGCTGTGAG 1132
 |||||
 Db 979 AGCAGAGCTGTGTACAAATTTAAGGGGCTTCAATGATGATGCAACCCATCCGCTGTGAG 1038
 QY 1133 AGCCTTATCTGAGATGATGATTAACCGCTGATGTGCTGCTGAGAACCTGCTGTGA 1192
 |||||
 Db 1039 AGCCTTATCTGAGATGATGATTAACCGCTGATGTGCTGCTGAGAACCTGCTGTGA 1098
 QY 1193 GAGACGAGCCCTTTACATCTTGTACCGGAGCATGAGCTGTGTACAGAGCTCTCGTTC 1252

AAK94129
 ID AAK94129 standard: cDNA, 1538 BP.
 AC AAK94129;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human full-length cDNA, SEQ ID NO: 2627.
 XX
 KM Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
 XX
 OS Homo sapiens.
 XX
 PN EPI130094-A2.
 XX
 PD 05-SEP-2001.
 XX
 PE 07-JUL-2000; 2000EP-0114089.
 XX
 PR 08-JUL-1999; 98JP-0194486.
 PR 11-JAN-2000; 2000JP-0118774.
 PR 02-MAY-2000; 2000JP-0183765.
 XX
 PA (HELI-) HELIX RES INST.
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Makamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 DR WPI; 2001-524255/58.
 DR P-PSDB; AAM93220.
 XX
 PT 830 Primers useful for synthesizing full length cDNA clones and their
 PT use in genetic manipulation -
 XX
 PS Claim 8; SEQ ID NO 2627; 1380bp + sequence listing; English.
 XX
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesised by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is a full length
 CC human cDNA of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.
 CC
 SQ Sequence 1538 BP; 387 A; 399 C; 350 G; 402 T; 0 other;
 Query Match 65.9%; Score 1534.8; DB 22; Length 1538;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1536; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 790 GTGAACGAGTGTGCCACCGAAGACCCCTGGTGCMAACCTGCGTCAACACCTAGCGCTCT 849
 DB 1 GTGAACGAGTGTGCCACCGAAGACCCCTGGTGCMAACCTGCGTCAACACCTAGCGCTCT 60
 QY 850 TTTCATCTGCGCTGTACCCAGATATGAACTTGAAGAAATGGCGTTCATTTGAGTAT 909
 DB 61 TTTCATCTGCGCTGTACCCAGATATGAACTTGAAGAAATGGCGTTCATTTGAGTAT 120
 QY 910 ATGAGCAGTGTGAGCTTCTGTAGTTCCTTGCACAAATGATGTGTGAACACCGCCGCG 969
 DB 121 ATGAGCAGTGTGAGCTTCTGTAGTTCCTTGCACAAATGATGTGTGAACACCGCCGCG 180
 QY 970 ACATATCTTGTCTCTGCGCTCCAGGCTACATCTGTGTGATGACCAACGAGCTGCCAA 1029
 DB 181 ACATATCTTGTCTCTGCGCTCCAGGCTACATCTGTGTGATGACCAACGAGCTGCCAA 240
 QY 1030 GACATCAACGAATGTGAGCAGAGAACCAACACCTGCAACCTGACAGACCTGTACAT 1089
 |||||||

DB 241 GACATCAACGAATGTGAGCAGAGAACCAACACCTGCAACCTGACAGACCTGTACAT 300
 QY 1090 TTACAAAGGGGCTTCAAAATGCATCGACCCCATCGGTGTGAGAGGCTTATGAGATC 1149
 DB 301 TTACAAAGGGGCTTCAAAATGCATCGACCCCATCGGTGTGAGAGGCTTATGAGATC 360
 QY 1150 AGTATTAACCGCTGTATGTCTCTGTGAGAACCTGTGTCAGAGACACGACCTTTACC 1209
 DB 361 AGTATTAACCGCTGTATGTCTCTGTGAGAACCTGTGTCAGAGACACGACCTTTACC 420
 QY 1210 ATCTTGTACCGGAGACGTGAGCTGTGTGACGACCTCGTTCCTCCGTGACATCTTCCA 1269
 DB 421 ATCTTGTACCGGAGACGTGAGCTGTGTGACGACCTCGTTCCTCCGTGACATCTTCCA 480
 QY 1270 ATGCAAGCAGACACCGCTACCTGTGGGGCTATTAATTTTCCAGATCAATCTGGGAAT 1329
 DB 481 ATGCAAGCAGACACCGCTACCTGTGGGGCTATTAATTTTCCAGATCAATCTGGGAAT 540
 QY 1330 GAGGCGAGAGAAATTTTACATGCGGCAACGCGGCCCATCAGTGCACCTGTGTATGACA 1389
 DB 541 GAGGCGAGAGAAATTTTACATGCGGCAACGCGGCCCATCAGTGCACCTGTGTATGACA 600
 QY 1390 CGCCCATCAAAAGGCGCCCGGAAATTCACAGCTGTGAAATGATCAGTCAACACT 1449
 DB 601 CGCCCATCAAAAGGCGCCCGGAAATTCACAGCTGTGAAATGATCAGTCAACACT 660
 QY 1450 GTCATCAACTTCAGAGGACGCTCGTGTATCGAGCTGTGAGATATATGTGTGCACTACCCA 1509
 DB 661 GTCATCAACTTCAGAGGACGCTCGTGTATCGAGCTGTGAGATATATGTGTGCACTACCCA 720
 QY 1510 TTCTGAGCTCGGGCTGTGAGCTTCGACGCTGCTCTCATTTGACCAAGGAGACAGAGA 1569
 DB 721 TTCTGAGCTCGGGCTGTGAGCTTCGACGCTGCTCTCATTTGACCAAGGAGACAGAGA 780
 QY 1570 AGAGGGAATTAACAAGAGAAATGAGAGGACACAGAGCTTAGCATTTCTGTGTAAGC 1629
 DB 781 AGAGGGAATTAACAAGAGAAATGAGAGGACACAGAGCTTAGCATTTCTGTGTAAGC 840
 QY 1630 TTTTCCCGGAAGATGACCCCGGCTTCTGACTCTCACTGTACTATTTGACAGACCTGTCA 1689
 DB 841 TTTTCCCGGAAGATGACCCCGGCTTCTGACTCTCACTGTACTATTTGACAGACCTGTCA 900
 QY 1690 CCCTGAGAGACTTGCACACCCCGGCTTCTGACTCTCACTGTACTATTTGACAGACCTGTCA 1749
 DB 901 CCCTGAGAGACTTGCACACCCCGGCTTCTGACTCTCACTGTACTATTTGACAGACCTGTCA 960
 QY 1750 CTCCCGTGTAGAGAAATGTTGTGGAATTTTCAAGGCTTCAGTTATTTCCACTATTTT 1809
 DB 961 CTCCCGTGTAGAGAAATGTTGTGGAATTTTCAAGGCTTCAGTTATTTCCACTATTTT 1020
 QY 1810 CAAGAAGAAATAGATTAGTTTGGGGGCTGTGAGTGTATGTTCAAAAGCTGTGAACACT 1869
 DB 1021 CAAGAAGAAATAGATTAGTTTGGGGGCTGTGAGTGTATGTTCAAAAGCTGTGAACACT 1080
 QY 1870 TGTGTGACCTTCTTCACTTCTTCACTTCTTCACTTCTTCACTTCTTCACTTCTTCACT 1929
 DB 1081 TGTGTGACCTTCTTCACTTCTTCACTTCTTCACTTCTTCACTTCTTCACTTCTTCACT 1140
 QY 1930 CCGGAGCTGTGCGGGGAAACCTGGAGTAGCTTGTGCTTTTTCGTAACAGAGAAG 1989
 DB 1141 CCGGAGCTGTGCGGGGAAACCTGGAGTAGCTTGTGCTTTTTCGTAACAGAGAAG 1200
 QY 1990 CTATGTAAACAAACACACAGAGATGCAAGGCTTTTGAAGATGTGTTCAAAACCATG 2049
 DB 1201 CTATGTAAACAAACACACAGAGATGCAAGGCTTTTGAAGATGTGTTCAAAACCATG 1260
 QY 2050 CCGTGTATTTTCAACCAATTAAGAGTTTCACTGCTTAAATTTGATAACGGTTTAA 2109
 DB 1261 CCGTGTATTTTCAACCAATTAAGAGTTTCACTGCTTAAATTTGATAACGGTTTAA 1320
 QY 2110 TTCTGCTTGTATTTTGAATTTTAAATATGCTGTAGATTCCTTCAAGAGGC 2169
 DB 1321 TTCTGCTTGTATTTTGAATTTTAAATATGCTGTAGATTCCTTCAAGAGGC 1380
 |||||||

DR P-PSDB: AAU75494.
XX New nucleic acid encoding human extracellular/epidermal growth factor,
PT useful for diagnosis and treatment of e.g. Marfan syndrome and wounds,
PT also related polypeptides -
XX
XX
PS Claim 4; Fig 1; 22pp; English.
XX The invention relates to a novel polynucleotide which is at
CC least 95% identical with a sequence (ATCC 97285) encoding mature human
CC extracellular protein-like/epidermal growth factor (EGF)-like protein,
CC EGF. Also included are the EGF EGF domains, a vector containing
CC the polynucleotide, a host cell containing the vector, anti-EGF
CC antibodies and antagonists of EGF. The polynucleotide is used for
CC recombinant production of EGF, in gene therapy, as hybridisation probes,
CC as antisense antagonists and for chromosome identification. The protein
CC is used to treat patients who require EGF, to identify specific
CC antagonists, used to treat conditions that require inhibition of EGF
CC (e.g. vascular smooth muscle cell proliferation, Marfan syndrome, wound
CC healing, neurological trauma, acquired immunodeficiency syndrome
CC (AIDS)-related dementia, ocular disorders, kidney disorders, liver
CC disorders, hair follicle growth promotion, burns, ulcers, corneal
CC incisions, corneal inflammation, neoplasms and psoriasis), to raise
CC specific antibodies and to characterise receptors. The present
CC sequence encodes EGF.
XX
SQ Sequence 1720 BP: 396 A; 515 C; 434 G; 372 T; 3 other:
Query Match 71.9%; Score 1674.2; DB 24; Length 1720;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1673; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 3 CCGGCGCTCTCCCGCTGCTCTCCAGACGCTGCGGCCCTCGAATAAACCC 62
DB 45 CCGGCGCTCTCCCGCTGCTCTCCAGACGCTGCGGCCCTCGAATAAACCC 104
QY 63 GCGAGCCCCGAGGCCCCAGAGAGGCGCAGCTGCCCCGAGCTCTCCGCGCCG 122
DB 105 GCGAGCCCCGAGGCCCCAGAGAGGCGCAGCTGCCCCGAGCTCTCCGCGCCG 164
QY 123 CGACCTTCTCTGCGCTTCCGATCTCCTCGCGGCGCTTGAGATCCAGGAATAA 182
DB 165 CGACCTTCTCTGCGCTTCCGATCTCCTCGCGGCGCTTGAGATCCAGGAATAA 224
QY 183 AAGGATCTCACTGTATACATTTGGCTGTCTGTCTTCCAAAGCCCTGGAAATCAGAGC 242
DB 225 AAGGATCTCACTGTATACATTTGGCTGTCTGTCTTCCAAAGCCCTGGAAATCAGAGC 284
QY 243 ACAGTGCACGATGAGCTTTGACCTGATGCGCAGTCAGAGAGTGTATGATATGATGA 302
DB 285 ACAGTGCACGATGAGCTTTGACCTGATGCGCAGTCAGAGAGTGTATGATATGATGA 344
QY 303 ATGCGGAACCATGCCCGAGGCGCTGCGAGAGACATGATGTGTTAAACCAAAATGGCG 362
DB 345 ATGCGGAACCATGCCCGAGGCGCTGCGAGAGACATGATGTGTTAAACCAAAATGGCG 404
QY 363 GATATTAATGATTTCCCGGACAAACCTGTGTATGAGAGGCCCTACTCAACCCCTACTC 422
DB 405 GATATTAATGATTTCCCGGACAAACCTGTGTATGAGAGGCCCTACTCAACCCCTACTC 464
QY 423 GACCCCTACTCAGGCTCGTATACCCAGAGCTGCCACCATCTCAGCTCAAACTATTC 482
DB 465 GACCCCTACTCAGGCTCGTATACCCAGAGCTGCCACCATCTCAGCTCAAACTATTC 524
QY 483 CAGCATCTCAGGCTCGTATATGCGCTTGATGATACAGATGATGAAGCAACCAATG 542
DB 525 CAGCATCTCAGGCTCGTATATGCGCTTGATGATACAGATGATGAAGCAACCAATG 564
QY 543 TGTGATGTGAGACAGTGTGCAACAGATTTCCACAGTGCACACCCACAGATCTGCAT 602
DB 585 TGTGATGTGAGACAGTGTGCAACAGATTTCCACAGTGCACACCCACAGATCTGCAT 644
QY 603 CAATAGTGAAGGCGGTACACCTGCTCTGACAGCAGGATATGCTTCTGGAAGGCCA 662

DB 645 CAATAGTGAAGGCGGTACACCTGCTCTGACAGCAGGATATGCTTCTGGAAGGCCA 704
QY 663 GTGCTTAGACATTTGATGATGATGCTGATATGTTACGACGACGCTCTGCGCAATTTCC 722
DB 705 GTGCTTAGACATTTGATGATGATGCTGATATGTTACGACGACGCTCTGCGCAATTTCC 764
QY 723 TGAATCCATTTCTGTATCATGCAACCTGTGTTTACCCATCAATGAGATGGAAGTCTTG 782
DB 765 TGAATCCATTTCTGTATCATGCAACCTGTGTTTACCCATCAATGAGATGGAAGTCTTG 824
QY 783 CCAAGATGTGACGAGTGTGTCACCCAGAAACCCCTGCTGCAAACTCGCTCAACCTTA 842
DB 825 CCAAGATGTGACGAGTGTGTCACCCAGAAACCCCTGCTGCAAACTCGCTCAACCTTA 884
QY 843 CGGCTCTTTCATCTGCGCCCTGAGACCCAGATATGATCACTTGAGAAATGAGGCTGATG 902
DB 885 CGGCTCTTTCATCTGCGCCCTGAGACCCAGATATGATCACTTGAGAAATGAGGCTGATG 944
QY 903 CAGTATATGAGACGAGTGCACCTTCTGAGTTCCTCTGCCAATGAGTGTGTAACCA 962
DB 945 CAGTATATGAGACGAGTGCACCTTCTGAGTTCCTCTGCCAATGAGTGTGTAACCA 1004
QY 963 GCCCGGACATATCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1022
DB 1005 GCCCGGACATATCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1064
QY 1023 CTGCGAAGACATCAAGATGAGATGAGACAGAGAACACAGTGCACCTGAGAGAGT 1082
DB 1065 CTGCGAAGACATCAAGATGAGATGAGACAGAGAACACAGTGCACCTGAGAGAGT 1124
QY 1083 CTCAATTTTACAAGGCGGCTTCAATATGATCGACCCATCGCTGTGAGAGGCTTATCT 1142
DB 1125 CTCAATTTTACAAGGCGGCTTCAATATGATCGACCCATCGCTGTGAGAGGCTTATCT 1184
QY 1143 GAGATACATGATTAACCGCTGTATGTCTCTGTAACACCTGGCTGCAAGACAGCC 1202
DB 1185 GAGATACATGATTAACCGCTGTATGTCTCTGTAACACCTGGCTGCAAGACAGCC 1244
QY 1203 CTTTACCATCTTGTACCGGAGCATGAGCTGTGTCAGAGAGCTGCTCCGCTGACAT 1262
DB 1245 CTTTACCATCTTGTACCGGAGCATGAGCTGTGTCAGAGAGCTGCTCCGCTGACAT 1304
QY 1263 CTTCCAAATGCAAGCCAGACCCGCTACCTGCGGCTTATACATTTTCCAGATCAATC 1322
DB 1305 CTTCCAAATGCAAGCCAGACCCGCTACCTGCGGCTTATACATTTTCCAGATCAATC 1364
QY 1323 TGGGATATGAGGCGAGAAATTTTACATGCGGCAAAAGGCGCCCATATAGTGCACCTG 1382
DB 1365 TGGGATATGAGGCGAGAAATTTTACATGCGGCAAAAGGCGCCCATATAGTGCACCTG 1424
QY 1383 GATGACAGCGCCCATCAAAAGGCGCCGGAATTCAGCTGAGCTTGAATGATCACTGT 1442
DB 1425 GATGACAGCGCCCATCAAAAGGCGCCGGAATTCAGCTGAGCTTGAATGATCACTGT 1484
QY 1443 CAACATGTCTATCAACTTCAGAGGAGCTCCGTGATCCGACTGCGGATATATGTGCGCA 1502
DB 1485 CAACATGTCTATCAACTTCAGAGGAGCTCCGTGATCCGACTGCGGATATATGTGCGCA 1544
QY 1503 GTACCCATTTCTGAGCTCTGGGCTGAGAGCTCCGAGGCTGCTCTCATTTGGACCAAGGA 1562
DB 1545 GTACCCATTTCTGAGCTCTGGGCTGAGAGCTCCGAGGCTGCTCTCATTTGGACCAAGGA 1604
QY 1563 CAGGAGAAAGAGGAAATTAACAGAGATGAGAGGAGACAGAGCTTTGGGCTTTCTCT 1622
DB 1605 CAGGAGAAAGAGGAAATTAACAGAGATGAGAGGAGACAGAGCTTTGGGCTTTCTCT 1664
QY 1623 CTGAACGTTTCCCGAAGAGTACGCCGACTTCTGACTCTGCACCTGTACTATTG 1678
DB 1665 CTGAACGTTTCCCGAAGAGTACGCCGACTTCTGACTCTGCACCTGTACTATTG 1720

RESULT 11

Qy	243	ICAGGACGACGAAATGGCTTTGACCTGGATGCGCCAGTCAAGGACAGTTTAAATATGATTA	302
Db	285	ACAGTGACGAAATGGCTTTGACCTGGATGCGCCAGTCAAGGACAGTTTAAATATGATTA	344
Qy	303	ATGCCGAATCCATCCCGAGGCGCTGCCGAGGAGACATGATGTGTGTTAAACAAATGGCGG	362
Db	345	ATGCCGAATCCATCCCGAGGCGCTGCCGAGGAGACATGATGTGTGTTAAACAAATGGCGG	404
Qy	363	GTAATTTATGATTTCCCGGAGCAAAACCTGTGTATACGAGGGGCCCTATCGAATCCCTACAC	422
Db	405	GTAATTTATGATTTCCCGGAGCAAAACCTGTGTATACGAGGGGCCCTATCGAATCCCTACAC	464
Qy	423	GACCCCTACTCGAGTCCGTAACCCAGCAGGTGCCCCACACTCTCAGCTCCAAATATCC	482
Db	465	GACCCCTACTCGAGTCCGTAACCCAGCAGGTGCCCCACACTCTCAGCTCCAAATATCC	524
Qy	483	CACGATCTCCAGGCGCTTATATATGCCCTTTGGATACAGATGGATGAAAGCAACCAATG	542
Db	525	CACGATCTCCAGGCGCTTATATATGCCCTTTGGATACAGATGGATGAAAGCAACCAATG	584
Qy	543	TGTGGATGTGGAGAGTGTCAACAGATATCCACACAGTGCAAACCCACCAATCTGCAT	602
Db	585	TGTGGATGTGGAGAGTGTCAACAGATATCCACACAGTGCAAACCCACCAATCTGCAT	644
Qy	603	CATACTGAAAGCGGGTACACCTGCTCCTGCACGACGAGATTTGGCTTCTGGAAGGCCA	662
Db	645	CATACTGAAAGCGGGTACACCTGCTCCTGCACGACGAGATTTGGCTTCTGGAAGGCCA	704
Qy	663	GAGCTTAGACATTGGATGAATGTGCTATATGTTACTGCGCAGCAGCTCTGCGCAATTTCC	722
Db	705	GAGCTTAGACATTGGATGAATGTGCTATATGTTACTGCGCAGCAGCTCTGCGCAATTTCC	764
Qy	723	TGGATCCTATTCTTGTATACGCAACCTGTTTTACCTCAATGAGAGATGGAAGTCTTG	782
Db	765	TGGATCCTATTCTTGTATACGCAACCTGTTTTACCTCAATGAGAGATGGAAGTCTTG	824
Qy	783	CCAAAGATGCAAGAGTGTGCCACCGAAGAACCCCGCGGCAACCTGGGTCAACACTTA	842
Db	825	CCAAAGATGCAAGAGTGTGCCACCGAAGAACCCCGCGGCAACCTGGGTCAACACTTA	884
Qy	843	CGGCTCTTTCATCTGCCGCTGTGACCCAGATATGAACTTTGAGGAAGATGGCGTTCA	902
Db	885	CGGCTCTTTCATCTGCCGCTGTGACCCAGATATGAACTTTGAGGAAGATGGCGTTCA	944
Qy	903	CAGTGAATGAGCAGATGTCACACTTCTCGAAGTCCCTGCCCAACATGAGTGTGAACCA	962
Db	945	CAGTGAATGAGCAGATGTCACACTTCTCGAAGTCCCTGCCCAACATGAGTGTGAACCA	1004
Qy	963	GCCCGGACATACTTGTCTCCTGCGCTCCACAGGCTACATCTGCTGGATGACCAACGAG	1022
Db	1005	GCCCGGACATACTTGTCTCCTGCGCTCCACAGGCTACATCTGCTGGATGACCAACGAG	1064
Qy	1023	CTGCCAAGACATCAACGAAATGTAGCACAGGAACACACGTCACACTGTGACAGACAGTG	1082
Db	1065	CTGCCAAGACATCAACGAAATGTAGCACAGGAACACACGTCACACTGTGACAGACAGTG	1124
Qy	1083	CTCAAAATTTCAAGGGGCTTCAAAATGATCAACCCATCCGCTGTGAGAGAGCTTATCT	1142
Db	1125	CTCAAAATTTCAAGGGGCTTCAAAATGATCAACCCATCCGCTGTGAGAGAGCTTATCT	1184
Qy	1143	GAGGATCAGTATTAACCGCTGTATGTCTCTGTGAACCCCTGGCTGCAGACACAGCC	1202
Db	1185	GAGGATCAGTATTAACCGCTGTATGTCTCTGTGAACCCCTGGCTGCAGACACAGCC	1244
Qy	1203	CTTTACATCTTGAACCGGGACATGAGAGGTGGTGAAGAGCGTCCGTTCCCGGTGACAT	1262
Db	1245	CTTTACATCTTGAACCGGGACATGAGAGGTGGTGAAGAGCGTCCGTTCCCGGTGACAT	1304
Qy	1263	CTTCCAATGCAAGCCAGACCGCGTACCTGAGGCGCTATTACATTTTCCAGATCAATC	1322
Db	1305	CTTCCAATGCAAGCCAGACCGCGTACCTGAGGCGCTATTACATTTTCCAGATCAATC	1364
Qy	1323	TGGGAATGAGGGCAGAGATTTTACATGCGGCAAAAGGGGCCCATAGTGCACACCTGCT	1382

Db	1365	TGGGAATGAGGGCAGAGAAATTTTACATGCGGCAACAGGGGCCCATCTAGTGGACCCCTGGT	1424
QY	1383	GATGACACGGCCCATCATCAAGGGGCCGGGAAATCCAGCTGGACTGTGGAAATGATCACATGT	1442
Db	1425	GATGACACGGCCCATCAAGGGGCCGGGAAATCCAGCTGGACTGTGGAAATGATCACATGT	1484
QY	1443	CAACACTGTATCACTCAACTTCAGAGGACAGCTCCGTATCCGACTGGCGAATATATGTGTGCA	1502
Db	1485	CAACACTGTATCACTCAACTTCAGAGGACAGCTCCGTATCCGACTGGCGAATATATGTGTGCA	1544
QY	1503	GTACCCATTTCTGAGCCTTCGGGCTGTGAGCTTCGACGCTGCTCTCATTTGGACCAAGGA	1562
Db	1545	GTACCCATTTCTGAGCCTTCGGGCTGTGAGCTTCGACGCTCTCATTTGGACCAAGGA	1604
QY	1563	CAGGAGAAAGAGGAAATTAACAGAGAGATGAGAGCGCACAGAGCTTAGGCATTTCTGT	1622
Db	1605	CAGGAGAAAGAGGAAATTAACAGAGAGATGAGAGCGCACAGAGCTTAGGCATTTCTGT	1666
QY	1623	CTGAAGCTTTCCCGAAGAGAGTCAAGCCCGACTTCCTGACTCTCAACTGTACTATTG	1678
Db	1665	CTGAAGCTTTCCCGAAGAGAGTCAAGCCCGACTTCCTGACTCTCAACTGTACTATTG	1720

Result ID	ABK13627	standard; cDNA; 1720 BP.
XX	ABK13627	
XX	23-APR-2002	(first entry)
XX	Human	cDNA encoding extracellular protein-like/EGF-like protein, EEE
XX	Human; ss: gene; extracellular protein-like protein; EGF-like;	
KW	protein; epidermal growth factor; EGF; ATCC 97285; gene therapy;	
KW	vascular smooth muscle cell proliferation; Marfan syndrome;	
KW	wound healing; neurological trauma; acquired immunodeficiency syndr	
KW	Alzheimer-related dementia; ocular disorder; kidney disorder; liver diso	
KW	hair follicle growth promotion; burn; ulcer; corneal incision;	
KW	corneal inflammation; neoplasm; psoriasis.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	211..1557
FT		/*tag= a
FT		/product= "EGF"
FT		/note= "this CDS minus the stop codon (211-2554) is
FT		specifically claimed in claim 5"
FT	sig_peptide	211..285
FT		/*tag= b
FT	mat_peptide	286..1554
FT		/*tag= c
FT		/label= Mature_EGF
FT		/note= "this region is specifically claimed in
FT		claim 4"
XX		
PN	US2001051358-A1.	
XX		
PD	13-DEC-2001.	
XX		
PF	25-MAR-1999;	99US-0275805.
XX		
PR	11-APR-1997;	97US-0839525.
PR	10-APR-1996;	96WO-US05247.
XX		
PA	(OLSE/) OLSEN H S.	
XX	(LIHH/) LI H.	
PI	Olsen HS, Li H;	
XX		
DR	WPI, 2002-121417/16.	

```

Db      1360 TGGGAATGAGGAGAGAAATTATTCATGCGGCAAAACGGGCCCAATCAGTGCACCCCTGCT 1419
QY      1383 GATGACACCCCATTAAGAGGCCCCGGGAAATTCACCTGACCTTGGAAATGATCAGCTGT 1442
Db      1420 GATGACACCCCATTAAGAGGCCCCGGGAAATTCACCTGACCTTGGAAATGATCAGCTGT 1479
QY      1443 CAACACTGTCATCACTTCAGAGAGAGCTCCGTCATCCGATCGGAGATATATGTGTGCA 1502
Db      1480 CACACTGTCATCACTTCAGAGAGAGCTCCGTCATCCGATCGGAGATATATGTGTGCA 1539
QY      1503 GTACCCATTCCTGAGCCTCGGCTGGAGCCTCCGACGCTGCTCATATGGACCAAGGGA 1562
Db      1540 GTACCCATTCCTGAGCCTCGGCTGGAGCCTCCGACGCTGCTCATATGGACCAAGGGA 1599
QY      1563 CAGGAGAAAGAGGAATATACAGAGAAATGAGAGCCACACAGACGTTAGCATTTCTGT 1622
Db      1600 CAGGAGAAAGAGGAATATACAGAGAAATGAGAGCCACACAGACGTTAGCATTTCTGT 1659
QY      1623 CTGAACGTTTCCCGAAGAGTCAAGCCCGACCTTCCTACTCTCACTGTACTATTCAGAG 1682
Db      1660 CTGAACGTTTCCCGAAGAGTCAAGCCCGACCTTCCTACTCTCACTGTACTATTCAGAG 1719
QY      1683 CCTGTACCCCTGACAGACTTCCACCCCAAGTTCCTATGATACAGTATCAAAAAGTAT 1742
Db      1720 CCTGTACCCCTGACAGACTTCCACCCCAAGTTCCTATGATACAGTATCAAAAAGTAT 1779
QY      1743 ATCATTCGTCCTCCGATGAGAAATGTTGGTGAATTTTCAAGGCTTCAGTTTATTTCCA 1802
Db      1780 ATCATTCGTCCTCCGATGAGAAATGTTGGTGAATTTTCAAGGCTTCAGTTTATTTCCA 1839
QY      1803 CTATTTTCAAGAAATAGATTAGTGTGGGGGCTGAGTCTATGTTTCAAGACTGTG 1862
Db      1840 CTATTTTCAAGAAATAGATTAGTGTGGGGGCTGAGTCTATGTTTCAAGACTGTG 1899
QY      1863 AACAGCTTGTCTCACTTCTTCACTCTTCCACTCTCTCTCTCACTGTGTACTGCTTGG 1922
Db      1900 AACAGCTTGTCTCACTTCTTCACTCTTCCACTCTCTCTCTCACTGTGTACTGCTTGG 1959
QY      1923 CAAGAGCCGGGAGCTGGGGGGGAAACCTGGAGTAGTATGCTTTGGCGTAGACA 1982
Db      1960 CAAGAGCCGGGAGCTGGGGGGGAAACCTGGAGTAGTATGCTTTGGCGTAGACA 2019
QY      1983 GAGAGGCTATGTAAACAAACACAGCAGAGATCGAAGGGTTTATAGAAATGTGTTCAA 2042
Db      2020 GAGAGGCTATGTAAACAAACACAGCAGAGATCGAAGGGTTTATAGAAATGTGTTCAA 2079
QY      2043 AACCATGCTGTGATTTTCAACATATAAAGATTTCAGTTGCTCTT 2089
Db      2080 AACCATGCTGTGATTTTCAACATATAAAGATTTCAGTTGCTCTT 2126

```

RESULT 9

AAV62432

ID AAV62432 standard; DNA; 1720 BP.

AC AAV62432;

DT 25-JAN-1999 (first entry)

XX Human EGF genomic DNA.

Extracellular/epidermal growth factor-like protein; EGF; human; liver;
 vascular smooth muscle cell proliferation; neurology; pathology; AIDS;
 dementia; ocular; cornea; inflammation; tumor cell; kidney;
 wound; epithelium; surgery; skin graft; ulcer; diabetic; trophic;
 Marfan syndrome; treatment; hair follicle; embryogenesis; neoplasia;
 epidermal cell; cancer; psoriasis; detection; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
 FH CDS 211..1557

```

FT      /*tag= a
FT      sig-peptide 211..285
FT      mat-peptide 286..1557
FT      /*tag= b
FT      /*tag= c
FT      /*product= "EGF"
FT      /note= "Extracellular/epidermal growth factor-like"

```

MO9846746-A1.

22-OCT-1998.

11-APR-1997; 97MO-US06020.

11-APR-1997; 97MO-US06020.

(HUMA-) HUMAN GENOME SCI INC.

LI H, Olsen HS;

WPI: 1998-568728/48.

P-PSDB; AAW79739.

PS New isolated extracellular/epidermal growth factor - used for
 regulating vascular smooth muscle cell proliferation, e.g. for
 enhancing neurological functions or treating neoplasia and other
 disorders.

PS Claim 1a: Fig 1A-D; 62pp; English.

XX This sequence encodes a novel human extracellular/epidermal growth
 factor-like protein, EGF. This protein can be used to regulate
 CC vascular smooth muscle cell proliferation and for restoration or
 CC enhancement of neurological functions diminished as a result of other
 CC damaging pathologies such as AIDS dementia. The protein can also be used
 CC to treat senile dementia, ocular disorders such as corneal inflammation,
 CC for targeting tumor cells, for treating kidney disorders, for liver
 CC regeneration or treating liver dysfunction, for treating wounds including
 CC all cutaneous wounds, corneal wounds, and injuries to the
 CC epithelial-lined hollow organs of the body or resulting from trauma such
 CC as burns, abrasions and cuts as well as from surgical procedures such as
 CC surgical incisions and skin grafting. The polypeptides can also be used
 CC for treating chronic conditions, such as chronic ulcers, diabetic ulcers,
 CC other non-healing (trophic) conditions, to treat Marfan syndrome, to
 CC promote hair follicular development, to stimulate growth and
 CC differentiation of various epidermal and epithelial cells in vivo and in
 CC vitro and to stimulate embryogenesis. Antagonists to EGF can be used to
 CC treat neoplasia such as cancers or tumors, skin disorders such as
 CC psoriasis or corneal inflammation. The products can also be used for
 CC identifying EGF receptors, detection, diagnosis and drug screening.

XX Sequence 1720 BP; 396 A; 515 C; 434 G; 372 T; 3 other;

Query Match 71.9%; Score 1674.2; DB 19; Length 1720;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1673; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

QY      3 CCGGGGCTCTCCCGCTGCTCTCCAGAGCTGCTGCGCCCTGTGGAATTAACACC 62
Db      45 CCGGGGCTCTCCCGCTGCTCTCCAGAGCTGCTGCGCCCTGTGGAATTAACACC 104
QY      63 GCGAGCCCGGAGGCGCAGAGAGGCGGACGTGCCCGAGCTCTCGGGGGTCCGCCCG 122
Db      105 GCGAGCCCGGAGGCGCAGAGAGGCGGACGTGCCCGAGCTCTCGGGGGTCCGCCCG 164
QY      123 CGAGCTTCTTCTCGCTTGCATCTCTCTCGGCGCTTGGACATGCCAGAAATAA 182
Db      165 CGAGCTTCTTCTCGCTTGCATCTCTCTCGGCGCTTGGACATGCCAGAAATAA 224
QY      183 AAGGATACACAGTATTCATCTGCGCTCTGCTCTTCCAGAGCCCTGGGAATGCACAGC 242
Db      225 AAGGATACACAGTATTCATCTGCGCTCTGCTCTTCCAGAGCCCTGGGAATGCACAGC 284

```


404 CGAGCTTCTTCGCGCTTGCATCTCTCTCGCGGCTTGGACATGCCAGAAATAA 463
183 AAGGATACACTGATGTACATTCGCTCTCTCTTCCAAAGCCCTGGAAATGCACAGC 242
464 AAGGATACACTGATGTACATTCGCTCTCTCTTCCAAAGCCCTGGAAATGCACAGC 523
243 ACAGTGCAGGAATGGCTTTGACCTGATCGCCAGTGCAGGAGATGTTTGAATTTATGA 302
524 ACAGTGCAGGAATGGCTTTGACCTGATCGCCAGTGCAGGAGATGTTTGAATTTATGA 583
303 ATGCCCAACATCCCGGAGGCTGCGCAGAGACATGATGTTTAAACCAAAATGGCGG 362
584 ATGCCCAACATCCCGGAGGCTGCGCAGAGACATGATGTTTAAACCAAAATGGCGG 643
363 GATTTATGACATTCGCGGAGCAAAACCTGATGATGAGGGGCTTACTGGAACCCCTACTC 422
644 GATTTATGACATTCGCGGAGCAAAACCTGATGATGAGGGGCTTACTGGAACCCCTACTC 703
423 GACCCCTACTAGTCCGATCCGACAGAGCTGCCCCACACACTCTCAGCTCCAAACTATCC 482
704 GACCCCTACTAGTCCGATCCGACAGAGCTGCCCCACACACTCTCAGCTCCAAACTATCC 763
483 CAGGATCTCAGGCTCTTATGATGCGCTTGGATACCAATGATGAAGAACCAATG 542
764 CAGGATCTCAGGCTCTTATGATGCGCTTGGATACCAATGATGAAGAACCAATG 823
543 TGTGATGTGAGAGATGTGCAACAGATTCACACAGTGCACACCCACAGATCTGAT 602
824 TGTGATGTGAGAGATGTGCAACAGATTCACACAGTGCACACCCACAGATCTGAT 883
603 CAATACTGAAGCGGCTACACCTGCTCTGCAACCGACGAGATATGGCTTGTGAAGGCA 662
884 CAATACTGAAGCGGCTACACCTGCTCTGCAACCGACGAGATATGGCTTGTGAAGGCA 943
663 GTGCTAGACATGATGAATGTGCTATGTTACTGACAGAGCTCTGTGGAAATGTTCC 722
944 GTGCTAGACATGATGAATGTGCTATGTTACTGACAGAGCTCTGTGGAAATGTTCC 1003
723 TGGATCTTATCTTGTACATGCAACCTGGTTTACCCTCAATGAGATGGAAGTCTTG 782
1004 TGGATCTTATCTTGTACATGCAACCTGGTTTACCCTCAATGAGATGGAAGTCTTG 1063
783 CCNAAGTGTGAAGAGATGTGCAACAGATTCACACAGTGCACACCCACAGATCTG 842
1064 CCNAAGTGTGAAGAGATGTGCAACAGATTCACACAGTGCACACCCACAGATCTG 1123
843 CGGCTCTTCACTGCGCTGTGACCCAGATATGATGAGAAAGTGGCGTTCATG 902
1124 CGGCTCTTCACTGCGCTGTGACCCAGATATGATGAGAAAGTGGCGTTCATG 1183
903 CAGTATATGAGAGATGTGCAACAGATTCACACAGTGCACACCCACAGATCTG 962
1184 CAGTATATGAGAGATGTGCAACAGATTCACACAGTGCACACCCACAGATCTG 1243
963 GCGCGGACATATCTTGTCTCTGCTCCCTCCAGGCTACATCTGCTGATGCAACCGAG 1022
1244 GCGCGGACATATCTTGTCTCTGCTCCCTCCAGGCTACATCTGCTGATGCAACCGAG 1303
1023 CTGCCAAGACATCAAGAAATGTGAGCAGAGAACACACAGTGCACACCGAGAGATG 1082
1304 CTGCCAAGACATCAAGAAATGTGAGCAGAGAACACACAGTGCACACCGAGAGATG 1363
1083 CTACAAATTTACAAAGGGGCTTCAAAATGCAATGCAACCCATCGCTGTGAGAGCCCTATCT 1142
1364 CTACAAATTTACAAAGGGGCTTCAAAATGCAATGCAACCCATCGCTGTGAGAGCCCTATCT 1423
1143 GAGGATCATGATTAACCGCTGTATGTCTGCTGAGAAACCTGCTGACAGACACCGC 1202
1424 GAGGATCATGATTAACCGCTGTATGTCTGCTGAGAAACCTGCTGACAGACACCGC 1483
1203 CTTTACCATCTTGTACCGGACATGAGAGTGTGTGAGAGAGCTCCGTTCCCGGTGACAT 1262
1484 CTTTACCATCTTGTACCGGACATGAGAGTGTGTGAGAGAGCTCCGTTCCCGGTGACAT 1543

1263 CTTCAAATGCAAGCCACGAGACCCGCTACCCCTGGGCTATTAACATTTTCAGATCAATC 1322
1544 CTTCAAATGCAAGCCACGAGACCCGCTACCCCTGGGCTATTAACATTTTCAGATCAATC 1603
1323 TGGGAATGAGGGCAGAGAAATTTTACATGCGGCAAAACGGGCCCATCAGTGCACCTGTG 1382
1604 TGGGAATGAGGGCAGAGAAATTTTACATGCGGCAAAACGGGCCCATCAGTGCACCTGTG 1663
1383 GATGACAGCGCCCATCAAAAGGGCCCCGGGAAATCCAGCTGACCTTGGAAATGATACGT 1442
1664 GATGACAGCGCCCATCAAAAGGGCCCCGGGAAATCCAGCTGACCTTGGAAATGATACGT 1723
1443 CAACACTGATCAACTTCAGAGGACGCTCGGTATCCGACTGGGATATATGTGCGCA 1502
1724 CAACACTGATCAACTTCAGAGGACGCTCGGTATCCGACTGGGATATATGTGCGCA 1783
1503 GTAACCATCTGAGCCCTCGGGCTGAGGCTCCGAGGCTGCTCTCATTTGGCAAGGGA 1562
1784 GTAACCATCTGAGCCCTCGGGCTGAGGCTCCGAGGCTGCTCTCATTTGGCAAGGGA 1800
1563 CAGGAGAGAGAGAAATTAACAGAGACATGAGAGCCACAGACGTTAAGCATTTCTG 1622
1801 -----CTG 1803
1623 CTGAAGCTTCCCGGAGAGTCAAGCCGACCTCTGACTGCTACCTGATATTCAGCA 1682
1804 CTGAAGCTTCCCGGAGAGTCAAGCCGACCTCTGACTGCTACCTGATATTCAGCA 1863
1683 CCGTGCACCTGCGAGACTTGCACACCCCACTTCTATGATGATATCAAAAGATAT 1742
1864 CCGTGCACCTGCGAGACTTGCACACCCCACTTCTATGATGATATCAAAAGATAT 1923
1743 ATCATGCTCCCTGATGAGAAATGTTGGAATTTCAAGGCTTACGTTATTTCCA 1802
1924 ATCATGCTCCCTGATGAGAAATGTTGGAATTTCAAGGCTTACGTTATTTCCA 1983
1803 CTATTTTCAAGAAATATAGATTAGTTTGCAGGGGTGAGTCTATGTTCAAAAGACTG 1862
1984 CTATTTTCAAGAAATATAGATTAGTTTGCAGGGGTGAGTCTATGTTCAAAAGACTG 2043
1863 AACAGCTGCTGACCTTCTTACCTCTTCCACTCTCTCTCACTGTTACTGCTTTG 1922
2044 AACAGCTGCTGACCTTCTTACCTCTTCCACTCTCTCTCACTGTTACTGCTTTG 2103
1923 CAAGAGCCGGGAGCTGGCGGGGAAACCTGGGAGTACTGTTGCTTTGGCAACCA 1982
2104 CAAGAGCCGGGAGCTGGCGGGGAAACCTGGGAGTACTGTTGCTTTGGCAACCA 2163
1983 GAGAGGCTATGTAACCAACACAGCAGATGCAAGGCTTTTGAAGATGTGTTCA 2042
2164 GAGAGGCTATGTAACCAACACAGCAGATGCAAGGCTTTTGAAGATGTGTTCA 2223
2043 AACCATGCTGCTATTTTCAACCATTAAGAAAGTTCAAGTGTCTTAAATTTGTATAC 2102
2224 AACCATGCTGCTGCTATTTTCAACCATTAAGAAAGTTCAAGTGTCTTAAATTTGTATAC 2283
2103 GGTATATCTGCTGCTGCTATTTTGAATTTTAAATAATATGCTAGATCTCTTC 2162
2284 GGTATATCTGCTGCTGCTATTTTGAATTTTAAATAATATGCTAGATCTCTTC 2343
2163 GAAAGGCTTTCAGACACATCTATGTTCTGCTTCCAAACCAAGTCTCTCCATTTT 2222
2344 GAAAGGCTTTCAGACACATCTATGTTCTGCTTCCAAACCAAGTCTCTCCATTTT 2403
2223 AGCCGAGTGTCTTCTTGAAGACCCCTTAATCTTGTCTTCTTGAATTTTACCAAT 2282
2404 AGCCGAGTGTCTTCTTGAAGACCCCTTAATCTTGTCTTCTTGAATTTTACCAAT 2463
2283 GATTTGGAATGCAAGGCTTCAAACTGATTAATTAATTTGAAGGA 2328
2464 GATTTGGAATGCAAGGCTTCAAACTGATTAATTAATTTGAAGGA 2509

PI Wood WI, Zhang Z;
XX WPI: 2001-050091/06.
DR P-PSDB: AAB31183.
XX
PT Isolated nucleic acid molecule encoding a PRO polypeptide which is a
PT transmembrane polypeptide is useful for gene therapy and identification
PT of related polypeptides -
XX
XX
PS Claim 2: Fig 9: 244pp: English.
CC The present sequence encodes a human secreted and transmembrane
CC polypeptide. The specification describes human polypeptides, designated
CC PRO156, PRO444, PRO183, PRO185, PRO210, PRO215, PRO217, PRO242, PRO288,
CC PRO355, PRO1361, PRO1308, PRO1183, PRO1272, PRO1419, PRO4999, PRO7170,
CC PRO248, PRO353, PRO1318, PRO1600, PRO9940, PRO533, PRO301, PRO187,
CC PRO337, PRO411, PRO4356, PRO246, PRO265, PRO941, PRO10096, PRO6003,
CC PRO6004, PRO350, PRO2650 and PRO6309. The biological activity of cells
CC can be modulated with agents that bind to these polypeptides, resulting
CC in the death of the cells. The polynucleotides encoding these
CC polypeptides are useful in the recombinant production of the
CC homologous sequences, or to map the gene. They may also be used for
CC analyzing genetic disorders, and to produce transgenic animals which are
CC useful for the development and screening of therapeutically useful
CC reagents. The polynucleotides can also be used in gene therapy e.g. to
CC replace a defective gene.
XX
XX Sequence 2609 BP: 622 A; 718 C; 622 G; 647 T; 0 other:
SQ
Query Match 99.88; Score 2322.8; DB 22; Length 2609;
Best Local Similarity 99.98; Pred. No. 0;
Matches 2324; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 3 CCGGGGCTTCCCGGCTGCTCCACGACTCGTGGGCCCCCTGGAATAAACACCC 62
DB 284 CCGGGGCTTCCCGGCTGCTCCACGACTCGTGGGCCCCCTGGAATAAACACCC 343
OY 63 GCGAGCCCCGAGGGCCGAGAGGAGGCGAGCTGCCCCAGCTCCCGGGGTCGCCGCCG 122
DB 344 GCGAGCCCCGAGGGCCGAGAGGAGGCGAGCTGCCCCAGCTCCCGGGGTCGCCGCCG 403
OY 123 CGAGCTTCTTCTGCGCTTGGCATCTCTCTCGCGGCTCTTGGACATGCCAGATAAA 182
DB 404 CGAGCTTCTTCTGCGCTTGGCATCTCTCTCGCGGCTCTTGGACATGCCAGATAAA 463
OY 183 AAGGATACACTGTTACATCTGCGCTCTGCTTCCCAAGCCCTGGGAATGACAGGC 242
DB 464 AAGGATACACTGTTACATCTGCGCTCTGCTTCCCAAGCCCTGGGAATGACAGGC 523
OY 243 ACAGTCACAGAAATGCTTGAACCTGATCGCCAGTCAGAGAGTGTTAATGATTGATGA 302
DB 524 ACAGTCACAGAAATGCTTGAACCTGATCGCCAGTCAGAGAGTGTTAATGATTGATGA 583
OY 303 ATGCCCAACCATCCCGAGGCTGCGAGAGAGATGATGTGTAAACCAAAATGGCGG 362
DB 584 ATGCCCAACCATCCCGAGGCTGCGAGAGAGATGATGTGTAAACCAAAATGGCGG 643
OY 363 GTATTATGATTTCCCGGGAACAAACCTGTATCGAGGGCCCTACTGAAACCCCTACAC 422
DB 644 GTATTATGATTTCCCGGGAACAAACCTGTATCGAGGGCCCTACTGAAACCCCTACAC 703
OY 423 GACCCCTACTAGTCCGATACCGAGAGTGCCTCCACATCTCAGCTCCAAATATATCC 482
DB 704 GACCCCTACTAGTCCGATACCGAGAGTGCCTCCACATCTCAGCTCCAAATATATCC 763
OY 483 CACGATCTCAGGCTCTTATATGCGCTTGGATACCATGATGATGAAGAACCAATG 542
DB 764 CACGATCTCAGGCTCTTATATGCGCTTGGATACCATGATGATGAAGAACCAATG 823
OY 543 TGTGATGTGAGAGAGTGTCAACAGATGCCACAGAGCAACCCACAGATGTGAT 602
DB 824 TGTGATGTGAGAGAGTGTCAACAGATGCCACAGAGCAACCCACAGATGTGAT 883

OY 603 CAATACGTAAGGGGTACACCTGCTCCGACACCGAGGATATGCTCTGGAAGGCCA 662
DB 884 CAATACGTAAGGGGTACACCTGCTCCGACACCGAGGATATGCTCTGGAAGGCCA 943
OY 663 GTGCTTACATGATGATGATGCTGATGCTTACTGCTCCAGACCTGTGCGAATGTCC 722
DB 944 GTGCTTACATGATGATGATGCTGATGCTTACTGCTCCAGACCTGTGCGAATGTCC 1003
OY 723 TGGATCCATCTCTGTATACATGCAACCTGTGTTTACCTTCATGAGATGGAAGTCTTG 782
DB 1004 TGGATCCATCTCTGTATACATGCAACCTGTGTTTACCTTCATGAGATGGAAGTCTTG 1063
OY 783 CCAAGATGTGAAGAGTGTGCGACCGAGAGAACCTGCGGCAAACTGGCTCAACACTTA 842
DB 1064 CCAAGATGTGAAGAGTGTGCGACCGAGAGAACCTGCGGCAAACTGGCTCAACACTTA 1123
OY 843 CGGCTCTTTCATCTGCGCTGTGACCCAGAGATGAACTTGAAGAGATGGCTTCATTTG 902
DB 1124 CGGCTCTTTCATCTGCGCTGTGACCCAGAGATGAACTTGAAGAGATGGCTTCATTTG 1183
OY 903 CAGTATATGAGACGAGTGCAGTCTGCTGATGCTTCCCTGCAACATGAGTGTGAACCA 962
DB 1184 CAGTATATGAGACGAGTGCAGTCTGCTGATGCTTCCCTGCAACATGAGTGTGAACCA 1243
OY 963 GCCCGGACATACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1022
DB 1244 GCCCGGACATACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1303
OY 1023 CTGCGAAGACATCAACGATGAGACAGACAGAACACACGTCGACAGAGAGTGG 1082
DB 1304 CTGCGAAGACATCAACGATGAGACAGACAGAACACACGTCGACAGAGAGTGG 1363
OY 1083 CTACAATTTAAAGGGGCTTCAATGATGCAACCCATCCCTGTGAGAGGCTTATCT 1142
DB 1364 CTACAATTTAAAGGGGCTTCAATGATGCAACCCATCCCTGTGAGAGGCTTATCT 1423
OY 1143 GAGATACAGTATTAACCCCTGATGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1202
DB 1424 GAGATACAGTATTAACCCCTGATGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1483
OY 1203 CTTTACCATCTTGTACCGGAGACAGTGTGTGACAGAGCTGCTGCTGCTGCTGCTGCTG 1262
DB 1484 CTTTACCATCTTGTACCGGAGACAGTGTGTGACAGAGCTGCTGCTGCTGCTGCTGCTG 1543
OY 1263 CTTTCAATATGACGACAGCCGCTTACCTGCGGCTTATCAATTTCCAGATCAATC 1322
DB 1544 CTTTCAATATGACGACAGCCGCTTACCTGCGGCTTATCAATTTCCAGATCAATC 1603
OY 1323 TGGGAATGAGGGGAGAGAAATTTTACATGCGGCAAAAGGCCCCCATGAGTGCACCTGGT 1382
DB 1604 TGGGAATGAGGGGAGAGAAATTTTACATGCGGCAAAAGGCCCCCATGAGTGCACCTGGT 1663
OY 1383 GATGACAGCGCCCATCAAAAGGGCCCCGGGAAATTCAGCTGAGTGGAAATGATCACTGT 1442
DB 1664 GATGACAGCGCCCATCAAAAGGGCCCCGGGAAATTCAGCTGAGTGGAAATGATCACTGT 1723
OY 1443 CAACACGTGTATCAACTTCAAGAGAGTCCGTGATCCGACTGCGGATATATGTGTCGGA 1502
DB 1724 CAACACGTGTATCAACTTCAAGAGAGTCCGTGATCCGACTGCGGATATATGTGTCGGA 1783
OY 1503 GTACCCATTTGAGAGCTTCGAGGCTGAGGCTCGGAGCTGAGCTCTCATTTGGACCAAGGA 1562
DB 1784 GTACCCATTTGAGAGCTTCGAGGCTGAGGCTCGGAGCTGAGCTCTCATTTGGACCAAGGA 1843
OY 1563 CAGGAGAGAGAGAGAAATTAACAGAGATGAGACGACACAGCTTATGAGCATTTCTGTG 1622
DB 1844 CAGGAGAGAGAGAGAAATTAACAGAGATGAGACGACACAGCTTATGAGCATTTCTGTG 1903
OY 1623 CTGAGAGTTCCTCCGAGAGAGTCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1682
DB 1904 CTGAGAGTTCCTCCGAGAGAGTCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1963

```

1083 CTACAAATTTACAAAGGGGGCTTCAATATGATGACCCCATCCGCTGTGAGAGACCTTATCT 1142
1364 CTACAAATTTACAAAGGGGGCTTCAATATGATGACCCCATCCGCTGTGAGAGACCTTATCT 1423
1143 GAGAGATGATGATTAACCGCTGTATGTCTCTGTGAGAAACCTGTGCTGACAGACAGCC 1202
1424 GAGAGATGATGATTAACCGCTGTATGTCTCTGTGAGAAACCTGTGCTGACAGACAGCC 1483
1203 CTTTACATCTTGTACCGGGGACATGAGACGTGTGTGACAGAGCTTCCTCCGCTGACAT 1262
1484 CTTTACATCTTGTGTACCGGGGACATGAGACGTGTGTGACAGAGCTTCCTCCGCTGACAT 1543
1263 CTTCAAATGCAAGCCAGACCCGCTACCCCTGGGGCCATTACATTTTCCAGATCAAAATC 1322
1544 CTTCCAAAATGCAAGCCAGACCCGCTACCCCTGGGGCCATTACATTTTCCAGATCAAAATC 1603
1323 TGGGAATGAGGCGACAGAAATTTTACATGCGGCAAAAGGGGCCCATCACTGCGCACCTGT 1382
1604 TGGGAATGAGGCGACAGAAATTTTACATGCGGCAAAAGGGGCCCATCACTGCGCACCTGT 1663
1383 GATGACACGGCCCATCAAAAGGGGGGAAATCCAGCTGTGACCTTGGAAATGATACATCT 1442
1664 GATGACACGGCCCATCAAAAGGGGGGAAATCCAGCTGTGACCTTGGAAATGATACATCT 1723
1443 CAACACTGTATCAACTTTCAGAGGACAGCTCCGTATCCGACTGCGGATATATGTGCGCA 1502
1724 CAACACTGTATCAACTTTCAGAGGACAGCTCCGTATCCGACTGCGGATATATGTGCGCA 1783
1503 GTACCAATCTGTAGCCCTCGGGGCTGTGAGACCTTCGACGCTCTTCATTTGGACCAAGGA 1562
1784 GTACCAATCTGTAGCCCTCGGGGCTGTGAGACCTTCGACGCTCTTCATTTGGACCAAGGA 1843
1563 CAGGAGAGAGAGAAATAACAGAGAGATGAGAGCGACAGACAGTGGATTCCTCTG 1622
1844 CAGGAGAGAGAGAAATAACAGAGAGATGAGAGCGACAGACAGTGGATTCCTCTG 1903
1623 CTGAACGTTTCCCGCGAAGAGTCAAGCCGACTTCCTACTCTCACTGTACTATTTGAGA 1682
1904 CTGAACGTTTCCCGCGAAGAGTCAAGCCGACTTCCTACTCTCACTGTACTATTTGAGA 1963
1683 CCTGTACCCCTGTGAGAGACTTGCACCCGACTTCCTATGATACAGTTATCAAAAAGTAT 1742
1964 CCTGTACCCCTGTGAGAGACTTGCACCCGACTTCCTATGATACAGTTATCAAAAAGTAT 2023
1743 ATCATTTGCTCCCTGTATAGAAATGTTGTGGAATTTTCAAGGGCTTCAGTTTATTTCCA 1802
2024 ATCATTTGCTCCCTGTATAGAAATGTTGTGGAATTTTCAAGGGCTTCAGTTTATTTCCA 2083
1803 CTATTTTCAAGAAATAGATTAAGTTTGGGGGGTCTGAGTCTATGTTTCAAGAGCTGTG 1862
2084 CTATTTTCAAGAAATAGATTAAGTTTGGGGGGTCTGAGTCTATGTTTCAAGAGCTGTG 2143
1863 AACAGCTTGTCTGCACTTCTTCACTCTTCTCTCACTCTCTCTCTCTCTCTCTCTCTCTG 1922
2144 AACAGCTTGTCTGCACTTCTTCACTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 2203
1923 CAAAGACCCGGAGCTGTGGGGGAAACCTCGGAGTACTGTTGCTTTTGGCGACACA 1982
2204 CAAAGACCCGGAGCTGTGGGGGAAACCTCGGAGTACTGTTGCTTTTGGCGACACA 2263
1983 GAGAAAGCTATGTAAACAAACACAGACAGATGAAAGGTTTTTGAAGATGTGTTTCA 2042
2264 GAGAAAGCTATGTAAACAAACACAGACAGATGAAAGGTTTTTGAAGATGTGTTTCA 2323
2043 AACCATGCTGTATTTTCAACCATTAAGAGAGTTGAGTGTCTTAAATTTGTATTAAC 2102
2324 AACCATGCTGTATTTTCAACCATTAAGAGAGTTGAGTGTCTTAAATTTGTATTAAC 2383
2103 GGTAAATCTGTCTGTCTTATTTAGTATTTTAAATAATNGTGTGAATTCCTTC 2162
2384 GGTAAATCTGTCTGTCTTATTTAGTATTTTAAATAATNGTGTGAATTCCTTC 2443
2163 GAAAGGCTTTCAGACACATGCTATGTTCTGTCTTCCCAACCCAGTCTCCTCTCATTTT 2222
```

```

Db 2444 GAAAGGCTTTCAGACACATGCTATGTTCTGTCTTCCAAACCCAGTCTCTCCATTTT 2503
QY 2223 AGCCCACTGTTTTCTTTGAGACCCCTTAATCTTCTTCTTTAGAAATTTTACCAATT 2282
Db 2504 AGCCCACTGTTTTCTTTGAGACCCCTTAATCTTCTTCTTTAGAAATTTTACCAATT 2563
QY 2283 GATTTGAAATGACAGAGTCTTCCAACTATTAATTTTGAAGA 2328
Db 2564 GATTTGAAATGACAGAGTCTTCCAACTATTAATTTTGAAGA 2609

RESULT 6
AAC86968
ID AAC86968 standard; cdna; 2609 BP.
XX
AC AAC86968;
XX
DT 20-APR-2001 (first entry)
XX
DE Nucleotide sequence of human polypeptide PRO210.
XX
KW Human; secreted protein; transmembrane protein; PRO196; PRO444; PRO183;
KW PRO185; PRO210; PRO215; PRO242; PRO288; PRO365; PRO1361; PRO1308;
KW PRO1183; PRO1272; PRO1419; PRO4999; PRO7170; PRO248; PRO353; PRO1318;
KW PRO1600; PRO9940; PRO533; PRO301; PRO187; PRO337; PRO1411; PRO4356;
KW PRO246; PRO265; PRO941; PRO10096; PRO6003; PRO6004; PRO350; PRO2630;
KW PRO6309; cell death; genetic disorder; transgenic animal; gene therapy;
ss.
XX
OS Homo sapiens.
XX
FH Key 450..1796
FT CDS /*tag= a
FT sig.peptide 450..524
FT /*tag= b
XX
PN MO200077037-A2.
XX
PD 21-DEC-2000.
XX
PF 22-MAY-2000; 2000MO-US14042.
XX
PR 15-JUN-1999; 99US-0139695.
PR 20-JUL-1999; 99US-0145070.
PR 26-JUL-1999; 99US-0145698.
PR 17-AUG-1999; 99US-0149396.
PR 01-SEP-1999; 99MO-US20111.
PR 08-SEP-1999; 99MO-US20594.
PR 15-SEP-1999; 99MO-US21090.
PR 15-SEP-1999; 99MO-US21547.
PR 30-NOV-1999; 99MO-US28313.
PR 01-DEC-1999; 99MO-US28301.
PR 02-DEC-1999; 99MO-US28565.
PR 07-DEC-1999; 99US-0169495.
PR 05-JAN-2000; 2000MO-US00219.
PR 18-FEB-2000; 2000MO-US04341.
PR 18-FEB-2000; 2000MO-US04342.
PR 22-FEB-2000; 2000MO-US04414.
PR 01-MAR-2000; 2000MO-US05601.
PR 02-MAR-2000; 2000MO-US05841.
PR 20-MAR-2000; 2000MO-US07377.
PR 30-MAR-2000; 2000MO-US08439.
PR 15-MAY-2000; 2000MO-US13358.
PR 17-MAY-2000; 2000MO-US13705.
XX
PA (GENE) GENENTECH INC.
XX
XX Ashkenazi AJ, Baker KP, Botstein DA, Desnoyers L, Eaton DL;
PI Ferrara N, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Gurney AL, Kijavica IJ, Mather JP, Napier MA, Pan J;
PI Paol NF, Roy MA, Stewart TA, Tamas D, Watanabe CK, Williams PM;
```

PR 06-MAR-2000; 2000US-186968P.
 PR 14-MAR-2000; 2000US-189320P.
 PR 14-MAR-2000; 2000US-189328P.
 PR 15-MAR-2000; 2000MO-US06884.
 PR 21-MAR-2000; 2000US-190828P.
 PR 21-MAR-2000; 2000US-191007P.
 PR 21-MAR-2000; 2000US-191048P.
 PR 21-MAR-2000; 2000US-191314P.
 PR 28-MAR-2000; 2000US-192655P.
 PR 29-MAR-2000; 2000US-193032P.
 PR 29-MAR-2000; 2000US-193053P.
 PR 30-MAR-2000; 2000MO-US08439.
 PR 04-APR-2000; 2000US-194449P.
 PR 04-APR-2000; 2000US-194647P.
 PR 11-APR-2000; 2000US-195975P.
 PR 11-APR-2000; 2000US-196000P.
 PR 11-APR-2000; 2000US-196187P.
 PR 11-APR-2000; 2000US-196590P.
 PR 11-APR-2000; 2000US-196820P.
 PR 18-APR-2000; 2000US-198121P.
 PR 18-APR-2000; 2000US-198585P.
 PR 25-APR-2000; 2000US-199397P.
 PR 25-APR-2000; 2000US-199550P.
 PR 03-MAY-2000; 2000US-201516P.
 PR 17-MAY-2000; 2000MO-US13705.
 PR 22-MAY-2000; 2000MO-US14042.
 PR 30-MAY-2000; 2000MO-US14941.
 PR 02-JUN-2000; 2000MO-US15264.
 PR 05-JUN-2000; 2000US-209832P.
 PR 28-JUL-2000; 2000MO-US20710.
 PR 22-AUG-2000; 2000US-0644848.
 PR 24-AUG-2000; 2000MO-US30952.
 PR 08-NOV-2000; 2000MO-US30952.
 PR 01-DEC-2000; 2000US-332678.
 PR 20-DEC-2000; 2000MO-US34956.
 PA (GETH) GENENTECH INC.
 PI Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
 PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
 PI WPI: 2001-602746/68.
 DR P-PSDB; AA029227.
 XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the
 PT presence of tumours, such as prostate and breast tumours, in mammals and
 PT to screen for modulators of the compounds -
 XX
 XX Claim 2; Fig 407; 774pp; English.
 CC Sequences AAS45925-AAS46231 represent DNA molecules encoding and PCR
 CC primers for PRO polypeptides of the invention. The sequences of the
 CC invention can be used to detect the presence of a tumour in a mammal by
 CC comparing the level of expression of a PRO polypeptide in a test sample
 CC of cells from the animal and a control sample of normal cells, whereby a
 CC higher level of expression in the test sample indicates the presence of a
 CC tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep,
 CC pigs, goats and rabbits but are preferably human. The polypeptides can be
 CC used to stimulate tumour necrosis factor (TNF) alpha release from human
 CC blood, when contacted with it. A specific polypeptide can be used to
 CC stimulate the proliferation or differentiation of chondrocyte cells. The
 CC PRO proteins can be used to determine the presence of tumours and also
 CC susceptibility to tumour development, particularly adrenal, lung, colon,
 CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
 CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
 CC can be used for genetic analysis of individuals with genetic disorders.
 XX
 XX Sequence 2609 BP; 622 A; 718 C; 622 G; 647 T; 0 other;

Query Match 99.8%; Score 2322.8; DB 22; Length 2609;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2324; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	3	CCCGGCGCTCTCCCGTCTCTCCACAGACTCGCTCGGCCCTCTGGAATAAACACCC	62
DB	284	CCCGGCGCTCTCCCGTCTCTCTCCACAGACTCGCTCGGCCCTCTGGAATAAACACCC	343
QY	63	GGGAGCCCCGAGGGCCAGAGAGGCGGACGTGCTCCGGGGGTCCGCCCG	122
DB	344	GGGAGCCCCGAGGGCCAGAGAGGCGGACGTGCTCCGGGGGTCCGCCCG	403
QY	123	CGAGCTTCTTCTCGGCTTGCATCTCTCTCGGCGCTTGGACATCCGGAATAA	182
DB	404	CGAGCTTCTTCTCGGCTTGCATCTCTCTCGGCGCTTGGACATCCGGAATAA	463
QY	183	AAGGATACACTCTTACATCTTGGCTCTCTCTCTTCCAAAGCCCTGGGAATGACAGCG	242
DB	464	AAGGATACACTCTTACATCTTGGCTCTCTCTCTTCCAAAGCCCTGGGAATGACAGCG	523
QY	243	ACAGTGCACGANTGGCTTTGACCTGGATGCGCAGTACAGACAGTGTATGATGATGA	302
DB	524	ACAGTGCACGANTGGCTTTGACCTGGATGCGCAGTACAGACAGTGTATGATGATGA	583
QY	303	ATGCGGAACCATCCCGGAGGCTCGCGGAGAGACATGATGTGTAAACCAAAATGGCGG	362
DB	584	ATGCGGAACCATCCCGGAGGCTCGCGGAGAGACATGATGTGTAAACCAAAATGGCGG	643
QY	363	GTAATTTATGATTCGCCGAGACAAACCCGTGTATGAGAGGCCCTACTGAAACCCCTACTC	422
DB	644	GTAATTTATGATTCGCCGAGACAAACCCGTGTATGAGAGGCCCTACTGAAACCCCTACTC	703
QY	423	GACCCCTTACTAGAGTCCGTACCCAGACAGCTGCCCCACACCTCTCAAGTCAACATATCC	482
DB	704	GACCCCTTACTAGAGTCCGTACCCAGACAGCTGCCCCACACCTCTCAAGTCAACATATCC	763
QY	483	CACGATCTCCAGGCTCTTATATGCGGCTTGGATGATGATGATGATGATGATGATGATG	542
DB	764	CACGATCTCCAGGCTCTTATATGCGGCTTGGATGATGATGATGATGATGATGATGATG	823
QY	543	TGTGATGTGAGAGAGTGTGCAACAGATTCACACAGTGCACACCCACACAGATCTGCAT	602
DB	824	TGTGATGTGAGAGAGTGTGCAACAGATTCACACAGTGCACACCCACACAGATCTGCAT	883
QY	603	CAATCTGGAAGGCGGTGACCTGCTCTGACACGACGATATTTGGCTTCTGGAAGGCCA	662
DB	884	CAATCTGGAAGGCGGTGACCTGCTCTGACACGACGATATTTGGCTTCTGGAAGGCCA	943
QY	663	GTGCTTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	722
DB	944	GTGCTTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1003
QY	723	TGATTCCTATTCCTTGTACATGCAACCCGTGTTTACCCCTCAATGAGATGAGATGCTTG	782
DB	1004	TGATTCCTATTCCTTGTACATGCAACCCGTGTTTACCCCTCAATGAGATGAGATGCTTG	1063
QY	783	CCAAGATGTGACAGAGTGTGCAACCCCTGCTGCAACCTGCTGCAACCTGCTGCAACCTG	842
DB	1064	CCAAGATGTGACAGAGTGTGCAACCCCTGCTGCAACCTGCTGCAACCTGCTGCAACCTG	1123
QY	843	CGGCTCTTATGATGCGGCTGTGACCCGAGATGATGATGATGATGATGATGATGATGATG	902
DB	1124	CGGCTCTTATGATGCGGCTGTGACCCGAGATGATGATGATGATGATGATGATGATGATG	1183
QY	903	CAGTATATGAGAGAGTGCACCTTCTGAGTTCCTGCAACATGAGTGTGTGAACCA	962
DB	1184	CAGTATATGAGAGAGTGCACCTTCTGAGTTCCTGCAACATGAGTGTGTGAACCA	1243
QY	963	GGCGGAGCATATCTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1022
DB	1244	GGCGGAGCATATCTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1303
QY	1023	CTGCAAGACATCAACGAATGTGAGACAGAGAACCAACGATGATGATGATGATGATGATG	1082
DB	1304	CTGCAAGACATCAACGAATGTGAGACAGAGAACCAACGATGATGATGATGATGATGATG	1363

Db	756	GTGTGATGTGGAGACGAGTGTGCACACGATTTCCACAGTGGCAACCCACACAGATCTGCAT	815
Qy	603	CAATACTGAAGCGGGGTACACCTGCTCTCTGCACCGAGCGGATTTGGCTTCTGGAGGCCA	662
Db	816	CAATACTGAAGGGGGGTACACCTGCTCTCTGCACCGAGCGGATTTGGCTTCTGGAGGCCA	875
Qy	653	GTGCTTAGACATGTGAGTAATGTGGCTATTTGGTTACTGGCAGAGAGCTGTGGCAATTTCTCC	722
Db	876	GTGCTTAGACATGTGAGTAATGTGGCTATTTGGTTACTGGCAGAGAGCTGTGGCAATTTCTCC	935
Qy	723	TGGATCTTATTTCTGTATACATGCAACCTGGTTTTACCTCAATGAGAGATGGAAGTCTTG	782
Db	936	TGGATCTTATTTCTGTATACATGCAACCTGGTTTTACCTCAATGAGAGATGGAAGTCTTG	995
Qy	783	CCAAAGTGTGAACGAGTGTGGCCACCGGAAGACCCCTGGGTGGCAACCTGGGTCAACACCTA	842
Db	996	CCAAAGTGTGAACGAGTGTGGCCACCGGAAGACCCCTGGGTGGCAACCTGGGTCAACACCTA	1055
Qy	843	CGGCTCTTTTCATCTGGCGCTGTGTGACCCAGAGATATGAACTTGAGGAAGATGGGCTTATTG	902
Db	1056	CGGCTCTTTTCATCTGGCGCTGTGTGACCCAGAGATATGAACTTGAGGAAGATGGGCTTATTG	1115
Qy	903	CAGTGTATTTGACAGAGTGTGACGCTTCTGTGAATTTCTGTGCCAACAATGATGTGTGAACCA	962
Db	1116	CAGTGTATTTGACAGAGTGTGACGCTTCTGTGAATTTCTGTGCCAACAATGATGTGTGAACCA	1175
Qy	963	GGCCGGCACATACTTCTGCTCCTGCTCCCTCCAGAGCTACATCCGCTGGATGACAAACGGAAG	1022
Db	1176	GGCCGGCACATACTTCTGCTCCTGCTCCCTCCAGAGCTACATCCGCTGGATGACAAACGGAAG	1235
Qy	1023	CTGCCAAAGACATCAACGATGTGAGCACAGGACACACGTCGAACCTGCAGACAGCTG	1082
Db	1236	CTGCCAAAGACATCAACGATGTGAGCACAGGACACACGTCGAACCTGCAGACAGCTG	1295
Qy	1083	CTACAAATTTTCAAGGGGGGCTTCAAAATGATGAGACCCATCCGCTGGAGAGGCTTATCT	1142
Db	1296	CTACAAATTTTCAAGGGGGGCTTCAAAATGATGAGACCCATCCGCTGGAGAGGCTTATCT	1355
Qy	1143	GAGGATCAGTGTAAACCGCTGTATGTGTCTGTGAGAAACCTGTGGTCAGAGACCAAGCC	1202
Db	1356	GAGGATCAGTGTAAACCGCTGTATGTGTCTGTGAGAAACCTGTGGTCAGAGACCAAGCC	1415
Qy	1203	CTTTACCATTGTGACCGGGGACATGAGAGCTGTGCAGAGAGCTCCGCTTCCGCTGCAT	1262
Db	1416	CTTTACCATTGTGACCGGGGACATGAGAGCTGTGCAGAGAGCTCCGCTTCCGCTGCAT	1475
Qy	1263	CTTCCAAATGCAAGCCAGCACGACCCGCTACCTGTGGGCGTATTACATTTTCCAGATCAATC	1322
Db	1476	CTTCCAAATGCAAGCCAGCACGACCCGCTACCTGTGGGCGCTATTACATTTTCCAGATCAATC	1535
Qy	1323	TGGGAATGAGGGCGAGAAATTTTACATCCGGCAAAACGGGCCCATAGTGCACCCGTGT	1382
Db	1536	TGGGAATGAGGGCGAGAAATTTTACATCCGGCAAAACGGGCCCATAGTGCACCCGTGT	1595
Qy	1383	GATGACACGCCCCCTCAAAAGGGCCCCCGGAAATCCAGCTGAGATTGGAATGATCACTGT	1442
Db	1596	GATGACACGCCCCCTCAAAAGGGCCCCCGGAAATCCAGCTGAGATTGGAATGATCACTGT	1655
Qy	1443	CAACACTGTATCACTTACAGAGGACGCTCCGTGATCCGATCCGATCGGATATATGTGTGCA	1502
Db	1656	CAACACTGTATCACTTACAGAGGAGGAGCTCCGTGATCCGATCCGATCGGATATATGTGTGCA	1715
Qy	1503	GTACCAATTTGAGAGCTGTGGGCTGGAGGCTCCGAGAGCTCCCTCAATTTGGCACCAAGGA	1562
Db	1716	GTACCAATTTTGTAGGCTGTGGGCTGGAGGCTCCGAGAGCTCCCTCAATTTGGCACCAAGGA	1775
Qy	1563	CAGGAGAAGAGAGAAATTAACAGAGAGATGAGAGCGACACAGACGTTAGCAATTTCTG	1622
Db	1776	CAGGAGAAGAGAGAAATTAACAGAGAGATGAGAGCGACACAGACGTTAGCAATTTCTG	1835
Qy	1623	CTGAAGCTTTCCCGAAGAGTGCAGCCCGGACTTCTCGACTCTCACCTGTACTATTTGAGA	1682
Db	1836	CTGAAGCTTTCCCGAAGAGTGCAGCCCGGACTTCTCGACTCTCACCTGTACTATTTGAGA	1895

QY	1683	CCCTGTACACCTGCAGAGACTTGTCCACCCCGACCTCTTATGATATGATATCAAAAAGTATT	1742
Db	1896	CCCTGTACACCTGCAGAGACTTGTCCACCCCGACCTCTTATGATATGATATCAAAAAGTATT	1955
QY	1743	ATCATTTCTCCCTCCGATATGAAAGATTGTGTGGATATTTTCAAGGCTTCAGTTTATTTCCA	1802
Db	1956	ATCATTTCTCCCTCCGATATGAAAGATTGTGTGGATATTTTCAAGGCTTCAGTTTATTTCCA	2015
QY	1803	CTATTTTCAAGAAAATATGATTAGATTGTTCGGGGGCTGTGACTATATGTTTCAAAACATGTG	1862
Db	2016	CTATTTTCAAGAAAATATGATTAGATTGTTCGGGGGCTGTGACTATATGTTTCAAAACATGTG	2075
QY	1863	AACAGCTTGCCTGCTACCTTCTTACCTCTTCCACCTCTTCCACCTCTTCTCTACCTGTACTGCTTTG	1922
Db	2076	AACAGCTTGCCTGCTACCTTCTTACCTCTTCCACCTCTTCCACCTCTTCTCTACCTGTACTGCTTTG	2135
QY	1923	CAAAAGACCCGGAGACTGTGGCGGGGAACCCCTGGAGATAGCTACTTTCCTTTTGGCTACACA	1982
Db	2136	CAAAAGACCCGGAGACTGTGGCGGGGAACCCCTGGAGATAGCTACTTTCCTTTTGGCTACACA	2195
QY	1983	GAGAAGGCTATGTATTAACAACACACAGATGCCAAGGCTTTTATGAGAAATGCTTTTCAA	2042
Db	2196	GAGAAGGCTATGTATTAACAACACACAGATGCCAAGGCTTTTATGAGAAATGCTTTTCAA	2255
QY	2043	AACCATGCGCTGTATTTTCAACCAATTAAGAAGTTTCAGTTGCTCCTTAAATTTGTATTAAC	2102
Db	2256	AACCATGCGCTGTATTTTCAACCAATTAAGAAGTTTCAGTTGCTCCTTAAATTTGTATTAAC	2315
QY	2103	GGTTTAATTCCTGTCTGTTCATTTTGAATATTTTAAAAAATATGTCTAGATAATTCCTTC	2162
Db	2316	GGTTTAATTCCTGTCTGTTCATTTTGAATATTTTAAAAAATATGTCTAGATAATTCCTTC	2375
QY	2163	GAAAGGCTTTAGACACATGCTATGTCTGTCTTCCAAACCGATCCCTCCATCATTTT	2222
Db	2376	GAAAGGCTTTAGACACATGCTATGTCTGTCTTCCAAACCGATCCCTCCATCATTTT	2435
QY	2223	AGCCAGATGTTTCTTTTGAAGACCCCTTAACTGTCTTCTTTAGAAATTTTACCACATT	2282
Db	2436	AGCCAGATGTTTCTTTTGAAGACCCCTTAACTGTCTTCTTTAGAAATTTTACCACATT	2495
QY	2283	GGATTGGATGCAGAGTCTCCAAATCATTAATTAATTTTGAAGAGA	2328
Db	2496	GGATTGGATGCAGAGTCTCCAAATCATTAATTAATTTTGAAGAGA	2541
RESULT 5			
AAS46128			
ID	AAS46128 standard; cDNA; 2609 BP.		
XX	AAS46128;		
AC			
XX	18-DEC-2001 (first entry)		
DE			
XX	Human DNA encoding PRO polypeptide sequence #204.		
XX	PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep; ss;		
KW	dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;		
KW	blood; chondrocyte cell; cell proliferation; cell differentiation; colon;		
KW	adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder;		
KW	PCR primer.		
OS	Homo sapiens.		
XX	WO200168848-A2.		
PN			
XX	20-SEP-2001.		
PD			
XX	28-FEB-2001; 2001WO-US06520.		
PF			
XX	01-MAR-2000; 2000MO-US05601.		
PR	02-MAR-2000; 2000MO-US05841.		
PR	03-MAR-2000; 2000US-187202P.		


```

OY 723 TGGATCTATTCTTGTACATGCAACCCGTGTTTACCCTCATGAGATGGAAGGCTTG 782
DB 740 TGGATCTATTCTTGTACATGCAACCCGTGTTTACCCTCATGAGATGGAAGGCTTG 799
OY 783 CCAAGATGTGAACGAGTGTCCACCGAAGAACCCCTGGCTGCAAACTCGTCACACCTA 842
DB 800 CCAAGATGTGAACGAGTGTCCACCGAAGAACCCCTGGCTGCAAACTCGTCACACCTA 859
OY 843 CCGCTCTTTTCATCTGCCGCTGTGACCCAGATATGAACTTGAAGAAATGGCGTTTCA 902
DB 860 CCGCTCTTTTCATCTGCCGCTGTGACCCAGATATGAACTTGAAGAAATGGCGTTTCA 919
OY 903 CAGTGTATGAGAGAGTGTGAGCTTCTGTGAGTTCCTGCGCAACATGATGTTGTAACA 962
DB 920 CAGTGTATGAGAGAGTGTGAGCTTCTGTGAGTTCCTGCGCAACATGATGTTGTAACA 979
OY 963 GCCCGGACATATCTTGTCTGTGCTGCTCCAGGCTACATCTGCTGATGATGACCAAG 1022
DB 980 GCCCGGACATATCTTGTCTGTGCTGCTCCAGGCTACATCTGCTGATGATGACCAAG 1039
OY 1023 CTGCCAAGATATGAGATGTGAGACAGAGAACCAACAGTGTGCAACTGTGACAGACGTG 1082
DB 1040 CTGCCAAGATATGAGATGTGAGACAGAGAACCAACAGTGTGCAACTGTGACAGACGTG 1099
OY 1083 CTACAAATTTACAAGGGGGCTTCAAAATGATGACACCCATCCGCTGTGAGAGACCTTATCT 1142
DB 1100 CTACAAATTTACAAGGGGGCTTCAAAATGATGACACCCATCCGCTGTGAGAGACCTTATCT 1159
OY 1143 GAGATCAGTGTATACCCGCTGTATGTCTCTGTGAGAACCCCTGGCTGACAGACCAAGC 1202
DB 1160 GAGATCAGTGTATACCCGCTGTATGTCTCTGTGAGAACCCCTGGCTGACAGACCAAGC 1219
OY 1203 CTTTACCATCTTGTACCGGGAGCATGAGCGTGTGTCAGAGACGCTCCGCTGAGCAT 1262
DB 1220 CTTTACCATCTTGTACCGGGAGCATGAGCGTGTGTCAGAGACGCTCCGCTGAGCAT 1279
OY 1263 CTTCCAAATGCAAGCCAGACGACCCGCTACCCGCGGCTATTATACATTTTCCAGATCAATC 1322
DB 1280 CTTCCAAATGCAAGCCAGACGACCCGCTACCCGCGGCTATTATACATTTTCCAGATCAATC 1339
OY 1323 TGGGAATGAGGGCAGAGATTTTACATGCGGCAAAAGGGGCCCATCACTGACCCCTGCT 1382
DB 1340 TGGGAATGAGGGCAGAGATTTTACATGCGGCAAAAGGGGCCCATCACTGACCCCTGCT 1399
OY 1383 GATGACACGCGCCATCAAAAGGGGCCGGAATTCAGCTGAGCTTGAATGATACATGT 1442
DB 1400 GATGACACGCGCCATCAAAAGGGGCCGGAATTCAGCTGAGCTTGAATGATACATGT 1459
OY 1443 CAACACTGTATCACTTACAGAGCAGCTCCGTGATCCGATCGAGATATATGTGTGCA 1502
DB 1460 CAACACTGTATCACTTACAGAGCAGCTCCGTGATCCGATCGAGATATATGTGTGCA 1519
OY 1503 GTACCATCTGTAGCTCGGGCTGTGAGCCTCCGACGCTGCTCATTTGGCACAAGGA 1562
DB 1520 GTACCATCTGTAGCTCGGGCTGTGAGCCTCCGACGCTGCTCATTTGGCACAAGGA 1579
OY 1563 CAGGAGAGAGAGAAATAACAGAGATGAGAGCAGACAGAGCTTAGGATTTCCG 1622
DB 1580 CAGGAGAGAGAGAAATAACAGAGATGAGAGCAGACAGAGCTTAGGATTTCCG 1639
OY 1623 CTGAACGTTTCCCGAAGAGTCAAGCCCGACTTCTGACTCTCACTTACTATTGCA 1682
DB 1640 CTGAACGTTTCCCGAAGAGTCAAGCCCGACTTCTGACTCTCACTTACTATTGCA 1699
OY 1683 CCGTGTACCTGTGAGAGCTTCCACCCGAGTTCCTATGATACAGTATCAAAAAGTAT 1742
DB 1700 CCGTGTACCTGTGAGAGCTTCCACCCGAGTTCCTATGATACAGTATCAAAAAGTAT 1759
OY 1743 ATCATGTCTCCCTGATAGAAAGATTTGTTGTAATTTCAAGGCTTACGTTTATTTCA 1802
DB 1760 ATCATGTCTCCCTGATAGAAAGATTTGTTGTAATTTCAAGGCTTACGTTTATTTCA 1819

```

```

OY 1803 CTATTTTCAAAAGAAATAGATTAGTTTGCAGGGGCTGTGAGTCTATGTTCAAGACGTG 1862
DB 1820 CTATTTTCAAAAGAAATAGATTAGTTTGCAGGGGCTGTGAGTCTATGTTCAAGACGTG 1879
OY 1863 AACAGCTTGTCTGCTTCTTCACTTCTTCCACTTCTTCTCTCTCTCTCTCTCTCTCTCT 1922
DB 1880 AACAGCTTGTCTGCTTCTTCACTTCTTCCACTTCTTCTCTCTCTCTCTCTCTCTCT 1939
OY 1923 CAAAGACCCGGGAGCTGGCGGGGAACCTGGGAGTACGTAAGTCTTTTTCGTAACNA 1982
DB 1940 CAAAGACCCGGGAGCTGGCGGGGAACCTGGGAGTACGTAAGTCTTTTTCGTAACNA 1999
OY 1983 GAGAAGCTATGTATTAACAAACACAGCAGATGCAAGGTTTATGAGAAATGTGTTCA 2042
DB 2000 GAGAAGCTATGTATTAACAAACACAGCAGATGCAAGGTTTATGAGAAATGTGTTCA 2059
OY 2043 AACCATGCTGTATTTTCAACATTAAGAAAGTTTCAAGTGTCTTAAATTTGATTAAC 2102
DB 2060 AACCATGCTGTATTTTCAACATTAAGAAAGTTTCAAGTGTCTTAAATTTGATTAAC 2119
OY 2103 GGTATATCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2162
DB 2120 GGTATATCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2179
OY 2163 GAAAGGCTTCAAGACATGCTATGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2222
DB 2180 GAAAGGCTTCAAGACATGCTATGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2239
OY 2223 AGCCAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2282
DB 2240 AGCCAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2299
OY 2283 GATTTGGAATGACAGAGCTTCAAACTGATTAATTAATTTGAAGAGA 2328
DB 2300 GATTTGGAATGACAGAGCTTCAAACTGATTAATTAATTTGAAGAGA 2345

```

```

RESULT 3
AAZ39892
ID AAZ39892 standard; cDNA; 2362 BP.
XX
AC AAZ39892;
XX
DT 21-FEB-2000 (first entry)
XX
DE cDNA sequence of the human secreted protein AK647.
XX
KW AK647; aortic tissue development; smooth muscle cell modulator; SCID;
KW nutritional supplement; vasculogenesis; embryonic development; infection;
KW cytokine activity; cell proliferation; cell differentiation; detect; HIV;
KW immune deficiency; haematopoiesis regulation; tissue regrowth; diagnose;
KW wound healing; restenosis; atherosclerosis; drug screen; ss.
XX
OS Homo sapiens.
XX
FH key Location/Qualifiers
FT CDS 185..1532
FT FT /*tag= a
FT FT /product= AK647
FT FT /note= "Secreted protein"
XX
PD MO960125-A2.
XX
PD 25-NOV-1999.
XX
PF 18-MAY-1999; 99WO-US10931.
XX
PR 19-MAY-1998; 98US-0081002.
XX
PR 21-MAY-1998; 98US-0083002.
XX
PA (GENY ) GENETICS INST INC.
XX
PI Jacobs K, McCoy JM, Racie L, Lavallie E, Treacy M, Evans C;

```

QY	1801	CACATATTTTCAAAGAAAATAGATTGTTGGGGGGCTGGAGTCATATGTTCAAAGCTG	1800
Db	1801	CACTATTTTCAAAGAAAATAGATTGTTGGGGGGCTGGAGTCATATGTTCAAAGCTG	1860
QY	1861	TGAACAGCTTGCCTGTACCTTCTTCACTCTTCTCTCACTGTGTTACTGCTT	1920
Db	1861	TGAACAGCTTGCCTGTACCTTCTTCACTCTTCTCTCACTGTGTTACTGCTT	1920
QY	1921	TGCAAAAGACCCGGGAGCTGGGGGGGAAACCTGGGAGTACCTAGTTGCTTTGGCGTACA	1980
Db	1921	TGCAAAAGACCCGGGAGCTGGGGGGGAAACCTGGGAGTACCTAGTTGCTTTGGCGTACA	1980
QY	1981	CAGAGAAGGCTATGTAAACAACCCAGCAGGATCGAAGGGTTTTAGAGAAATGTCTTC	2040
Db	1981	CAGAGAAGGCTATGTAAACAACCCAGCAGGATCGAAGGGTTTTAGAGAAATGTCTTC	2040
QY	2041	AAAAACATGTCCTGTATTTTCAACCATAAAGAAAGTTTCAGTTGCTCTTAAATTTGTATA	2100
Db	2041	AAAAACATGTCCTGTATTTTCAACCATAAAGAAAGTTTCAGTTGCTCTTAAATTTGTATA	2100
QY	2101	ACGGTTAAATTCCTGTCTTGTTCATTTTGTAGTATTTTAAATAATNGTGTAGAAATTCCT	2160
Db	2101	ACGGTTAAATTCCTGTCTTGTTCATTTTGTAGTATTTTAAATAATNGTGTAGAAATTCCT	2160
QY	2161	TCGAAAAGGCTTCAGACATGTCATGTCTTCTTCCCAAAACCCAGTCCCTCCATTT	2220
Db	2161	TCGAAAAGGCTTCAGACATGTCATGTCTTCTTCCCAAAACCCAGTCCCTCCATTT	2220
QY	2221	TTAGCCCAAGTGTCTTCTTGTGAGAACCCCTTAATCTGTCTTCTTTAGAAATTTTACC	2280
Db	2221	TTAGCCCAAGTGTCTTCTTGTGAGAACCCCTTAATCTGTCTTCTTTAGAAATTTTACC	2280
QY	2281	TTGGATTGGAATCGAGAGGCTCCAAACTGATTAATATTTGAAGAGA	2328
Db	2281	TTGGATTGGAATCGAGAGGCTCCAAACTGATTAATATTTGAAGAGA	2328
RESULT 2			
AA07567	AA07567	standard; cDNA; 2362 BP.	
AC	AA07567;		
DT	21-JUN-1999	(first entry)	
XX	Homo sapiens fetal kidney clone AK647	secreted protein gene.	
XX	Secreted protein; fetal kidney; ds.		
XX	Homo sapiens.		
XX	Key	Location/Qualifiers	
FT	CDS	186..1532	
FT		/*tag= a	
FT		/product= "secreted protein"	
PN	W09900405-A1.		
PD	07-JAN-1999.		
PF	29-JUN-1998;	98WO-US13530.	
PR	30-JUN-1997;	97US-0885610.	
PA	(GENY) GENETICS INST INC.		
PI	Agostino M, Evans C, Jacobs K, Lavallic ER, McCoy JM,		
PI	Merberg D, Racie LA, Treacy M;		
DR	WPI: 1999-095671/08.		
P-PSDB:	AA095709.		

Query Match	Best Local Similarity	Score 2326:	DB 20:	Length 2362:
Matches 2326:	Conservative	0:	Mismatches	0:
			Indels	Gaps
3	CCCGGCGCTCCCGCTGCTCTCCACAGACGCGCTGGCCCTCTGGAAATAAACACACC	62		
20	CCCGGCGCTCCCGCTGCTCTCCACAGACGCTGGCCCTCTGGAAATAAACACACC	79		
63	CGGAGCCCGGAGGCGCCAGAGAGGCGACGTCGCCAGCTCTCCGGGGTCCCGCCG	122		
80	CGGAGCCCGGAGGCGCCAGAGAGGCGACGTCGCCAGCTCTCCGGGGTCCCGCCG	139		
123	CGAGCTTCTCTGCGCTCTGCGATCTCTCTCCGCGGCTCTGGACATGCGAGGAATAA	182		
140	CGAGCTTCTCTGCGCTCTGCGATCTCTCTCCGCGGCTCTGGACATGCGAGGAATAA	199		
183	AAGGATCTCACTGTTACATCTTGGCTCTCTGTTTCCAAAGCCCTGGGAATGACAGGC	242		
200	AAGGATCTCACTGTTACATCTTGGCTCTCTGTTTCCAAAGCCCTGGGAATGACAGGC	259		
243	ACAGGCGACGAGTGGCTTGGACGCGAGTGGCGACGACGAGAGGTTAGATATTGATGA	302		
260	ACAGGCGACGAGTGGCTTGGACGCGAGTGGCGACGACGAGAGGTTAGATATTGATGA	319		
303	ATGCGGACCACTCCCGCGAGGCGCTCGGAGGAGACATGATGTGTATTAACCAAAATGGCG	362		
320	ATGCGGACCACTCCCGCGAGGCGCTCGGAGGAGACATGATGTGTATTAACCAAAATGGCG	379		
363	GTATTATATGCAATCCCGGACAAACCTGTGTATGAGGCGCTTACTGAAACCCCTACTC	422		
380	GTATTATATGCAATCCCGGACAAACCTGTGTATGAGGCGCTTACTGAAACCCCTACTC	439		
423	GACCCCTTACGAGTGGCTGTCACGAGAGCGTGGCCACACTTCAGTCCAAACTATCC	482		
440	GACCCCTTACGAGTGGCTGTCACGAGAGCGTGGCCACACTTCAGTCCAAACTATCC	499		
483	CAGCATCTCCAGGCGCTTATATATCCGTTTGGATTACAGATGAGTGAAGAACCAACATG	542		
500	CAGCATCTCCAGGCGCTTATATATCCGTTTGGATTACAGATGAGTGAAGAACCAACATG	559		
543	TGTGGATGTGGAGAGTGTGCAAGATTCACAGTGCAAACCCCAACCAAGATCTCAT	602		
560	TGTGGATGTGGAGAGTGTGCAAGATTCACAGTGCAAACCCCAACCAAGATCTCAT	619		
603	CAATATGAGAGGCGGAGTACACTTCTCTCGACCCGAGAGATATGGCTTGTGGAAGGCA	662		
620	CAATATGAGAGGCGGAGTACACTTCTCTCGACCCGAGAGATATGGCTTGTGGAAGGCA	679		
663	GTGCTTACATTTATATATGTGCTTATGTTACTGCGACGACGCTGTGTGGAATGTTC	722		
680	GTGCTTACATTTATATATGTGCTTATGTTACTGCGACGACGCTGTGTGGAATGTTC	739		

DR P-PSDB; AAY54989.

XX Novel human polypeptides for treatment of, e.g. arteriosclerosis and
PT myoma -

PS Claim 4, 5; Page 76-80; 87pp; Japanese.

XX This sequence encodes the human A55 protein of the invention. The protein
CC can be used for the treatment of diseases due to abnormal proliferation
CC of smooth muscle. The polypeptides can be used according their inhibition
CC of the proliferation of vascular smooth muscle cells, particularly in
CC treating arteriosclerosis or re-narrowing by vascular endothelial
CC thickening after percutaneous transluminal coronary angioplasty (PTCA),
CC or myoma, haemopoietic cell-regulatory activity, cytokine activity,
CC tissue generation/repairation activity, actin/inhibin activity, taxis
CC and chemotaxis activity, blood coagulation/thrombotic activity,
CC receptor/ligand activity, cachectin/tumour metastasis inhibiting activity;
CC Note: The open reading frame of this sequence, as well as the mature
CC protein coding sequence are specifically claimed in claim 4.

XX Sequence 2328 BP; 560 A; 646 C; 541 G; 581 T; 0 other:

Query Match 100.0%; Score 2328; DB 21; Length 2328;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCCGGCGCTCCCGGTCTCTCCAGCAGTCGCTGGCCCTCTGGAATAAACAAC 60
DB 1 GACCCGGCGCTCCCGGTCTCTCCAGCAGTCGCTGGCCCTCTGGAATAAACAAC 60
QY 61 CCGCGAGCCCGGAGGCGCCAGAGAGGCGAGCGCGGAGCTCTCGGGGGTCCCGCC 120
DB 61 CCGCGAGCCCGGAGGCGCCAGAGAGGCGAGCGCGGAGCTCTCGGGGGTCCCGCC 120
QY 121 CCGGAGCTTCTCTCGCTTCGCTTCCTCTCCAGCAGTCGCTGGAGATGCGAGATA 180
DB 121 CCGGAGCTTCTCTCGCTTCGCTTCCTCTCCAGCAGTCGCTGGAGATGCGAGATA 180
QY 181 AAAAGATACACTGTATACATCTGTGCTCTCTCCAGCAGTCGCTGGAGATGCGAG 240
DB 181 AAAAGATACACTGTATACATCTGTGCTCTCTCCAGCAGTCGCTGGAGATGCGAG 240
QY 241 GCACAGTCGACAGATGCTTGAACCTGGATGCGAGATGCTTGAATATATATAT 300
DB 241 GCACAGTCGACAGATGCTTGAACCTGGATGCGAGATGCTTGAATATATATATAT 300
QY 301 GAATGCCGAACATCCCGGAGGCTGCGAGAGAGATGATGTGTAAACCAAAATGCG 360
DB 301 GAATGCCGAACATCCCGGAGGCTGCGAGAGAGATGATGTGTAAACCAAAATGCG 360
QY 361 GGGTATTTATGATTCCTCCGAGCAAAACCTGTGTATGAGGCGCTTACTCGAACCTTAC 420
DB 361 GGGTATTTATGATTCCTCCGAGCAAAACCTGTGTATGAGGCGCTTACTCGAACCTTAC 420
QY 421 TCGACCCCTCTACAGGTCCGTACCAAGACAGTCCGACACAGTCTGAGCTCAAACTAT 480
DB 421 TCGACCCCTCTACAGGTCCGTACCAAGACAGTCCGACACAGTCTGAGCTCAAACTAT 480
QY 481 CCCAGATCTCAGGCGCTTATATGCGCTTGTGATACAGATGATGAAGCAACCA 540
DB 481 CCCAGATCTCAGGCGCTTATATGCGCTTGTGATACAGATGATGAAGCAACCA 540
QY 541 TGTGTGATGATGACAGATGTCGAACAGATTCACACAGTGAACCCACAGATGTGC 600
DB 541 TGTGTGATGATGACAGATGTCGAACAGATTCACACAGTGAACCCACAGATGTGC 600
QY 601 ATCAATACAGAGGCGGTACACTGCTCTGCAACGAGATATGGCTTCTGGAAGC 660
DB 601 ATCAATACAGAGGCGGTACACTGCTCTGCAACGAGATATGGCTTCTGGAAGC 660
QY 661 CAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 661 CAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720

DB 661 CAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 721 CTTGATTCCTATTTTGTATGCAACCTGGTTTACCCTCAATGAGATGGAAGTCT 780
DB 721 CTTGATTCCTATTTTGTATGCAACCTGGTTTACCCTCAATGAGATGGAAGTCT 780
QY 781 TGGCAATGATGAGAGAGTGTGCGACGAGAACCCCTGGTGCACAAACCGGCGCAAC 840
DB 781 TGGCAATGATGAGAGAGTGTGCGACGAGAACCCCTGGTGCACAAACCGGCGCAAC 840
QY 841 TAGGCTCTTTCATCTGCGCTGTGACCCAGATATGAACCTTGAAGAGATGGCTTAT 900
DB 841 TAGGCTCTTTCATCTGCGCTGTGACCCAGATATGAACCTTGAAGAGATGGCTTAT 900
QY 901 TGCAGTATATGACAGAGTGTGAGTCTGTGATTCCTTGTCCAAATGATGATGTAAC 960
DB 901 TGCAGTATATGAGAGAGTGTGAGTCTGTGATTCCTTGTCCAAATGATGATGTAAC 960
QY 961 CAGCCGGGACATCTCTGCTGCTGCGCTCCAGGCTCAATCTGCTGATGCAACCGA 1020
DB 961 CAGCCGGGACATCTCTGCTGCTGCGCTCCAGGCTCAATCTGCTGATGCAACCGA 1020
QY 1021 AGCTGCCAAGACATCAAGATGTGAGCAGAGAACCAACGTCGACAGCAGAG 1080
DB 1021 AGCTGCCAAGACATCAAGATGTGAGCAGAGAACCAACGTCGACAGCAGAG 1080
QY 1081 TGTACAAATTTACAAAGGGGCTTCAATGATGATGATGATGATGATGATGAT 1140
DB 1081 TGTACAAATTTACAAAGGGGCTTCAATGATGATGATGATGATGATGATGAT 1140
QY 1141 CTGAGATAGATGATTAACCGCTGTATGCTGCTGTGAGAACCTGTGTCAGAGAC 1200
DB 1141 CTGAGATAGATGATTAACCGCTGTATGCTGCTGTGAGAACCTGTGTCAGAGAC 1200
QY 1201 CCTTTACATCTTGTACCGGAGATGAGAGTGTGTGAGAGCGTCCCTCCGCTGAC 1260
DB 1201 CCTTTACATCTTGTACCGGAGATGAGAGTGTGTGAGAGCGTCCCTCCGCTGAC 1260
QY 1261 ATCTTCCAAATGCAAGCCACACACCGCTACCTGGGGCTTATTAATTTTCCAGATCAA 1320
DB 1261 ATCTTCCAAATGCAAGCCACACACCGCTACCTGGGGCTTATTAATTTTCCAGATCAA 1320
QY 1321 TCTGGAGATGAGGCGAGAAATTTTACATGCGGCAAAAGGGGCCCATCAGTCCACCTG 1380
DB 1321 TCTGGAGATGAGGCGAGAAATTTTACATGCGGCAAAAGGGGCCCATCAGTCCACCTG 1380
QY 1381 GTGATGACAGGCGCCATCAAAAGGGGCCGGAATTCAGCTGAGATGGAATGATCAGCT 1440
DB 1381 GTGATGACAGGCGCCATCAAAAGGGGCCGGAATTCAGCTGAGATGGAATGATCAGCT 1440
QY 1441 GTCAACACTGTCAATCACTTGAAGAGAGCTCCGATGCCAGTCCGGATATATGTGTG 1500
DB 1441 GTCAACACTGTCAATCACTTGAAGAGAGCTCCGATGCCAGTCCGGATATATGTGTG 1500
QY 1501 CAGTACCATTTCTAGGCTCGGGCTGAGAGCTCCGACGCTGCTCATTTGGCACCACAA 1560
DB 1501 CAGTACCATTTCTAGGCTCGGGCTGAGAGCTCCGACGCTGCTCATTTGGCACCACAA 1560
QY 1561 GACAGGAGAGAGAGAAATTAACAGAGAGATGAAGAGACACAGACGTTAGGCAATTTCC 1620
DB 1561 GACAGGAGAGAGAGAAATTAACAGAGAGATGAAGAGACACAGACGTTAGGCAATTTCC 1620
QY 1621 TGTGAGAGCTTTCCCGGAGAGTGAAGCCGACCTCTGACCTCACTGCTATATGGA 1680
DB 1621 TGTGAGAGCTTTCCCGGAGAGTGAAGCCGACCTCTGACCTCACTGCTATATGGA 1680
QY 1681 GACCTGTACACCTTCAGAGACTTGCACACCCAGTTCATATGATATGAGTATTAATTTTC 1740
DB 1681 GACCTGTACACCTTCAGAGACTTGCACACCCAGTTCATATGATATGAGTATTAATTTTC 1740
QY 1741 TTATCATTTGCTCCCTGATAGAGATGTTGTGTAATTTTCAAGGCTTCAGTTATTTTC 1800
DB 1741 TTATCATTTGCTCCCTGATAGAGATGTTGTGTAATTTTCAAGGCTTCAGTTATTTTC 1800

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 08:40:41 ; Search time 518.747 Seconds
(without alignments)
10106.379 Million cell updates/sec

Title: US-09-674-379a-12

Perfect score: 2328
Sequence: 1 gaccgcgcgcctcccccgtg.....tgattaatatttcgaagaga 2328

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N.Geneseq_101002:*

- 1: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2328	100.0	2328	21	AAZ40027
2	2326	99.9	2326	20	AAZ40027
3	2326	99.9	2326	21	AAZ39892
4	2324.4	99.8	2550	20	AAZ05359
5	2323.8	99.8	2609	22	AAZ46128
6	2322.8	99.8	2609	22	AAZ46128
7	2112.8	90.8	2509	20	AAZ37670
8	2085.4	89.6	2126	22	AAZ94505
9	1674.2	71.9	1720	19	AAV62432

10	1674.2	71.9	1720	24	ABK13627	Human cDNA encoding
11	1534.8	65.9	1538	22	AAK94129	Human full-length
12	1480.8	63.6	2233	21	AAZ39384	Smooth muscle prol
13	1480.8	63.6	2233	21	AAZ40029	Full length mouse
14	1449.8	62.3	2429	21	AAZ39387	Smooth muscle prol
15	1449.8	62.3	2429	21	AAZ40031	Full length mouse
16	1347	57.9	1347	24	ABL52528	Human up50 cDNA se
17	1222.2	52.5	1228	18	AAZ89380	Human extracellular
18	1113.6	47.8	1344	21	AAZ39383	Smooth muscle prol
19	1098.6	47.2	1383	21	AAZ39386	Smooth muscle prol
20	1056.2	45.4	1269	21	AAZ39385	Smooth muscle prol
21	1056.2	45.4	1269	21	AAZ39388	Smooth muscle prol
22	898	38.6	1254	21	AAZ61672	CDNA encoding rat
23	898	38.6	1254	22	AAZ92195	Skin cell cDNA, SE
24	896	38.5	1250	24	ABZ34757	Rat cDNA isolated
25	813.6	34.9	853	22	AAK31556	Human cDNA 5'-end
26	813.6	34.9	853	22	AAK31559	Human cDNA clone r
27	776.4	33.4	778	20	AAZ37673	Human PRO protein
28	636.4	27.3	680	21	AAZ44245	Human secreted exp
29	556	23.9	670	22	AAK92195	Human cDNA 5'-end
30	556	23.9	670	22	AAK93463	Human cDNA clone r
31	493	21.2	587	22	AAK92416	Human cDNA 3'-end
32	448.4	19.3	519	18	AAZ88061	Partial cDNA clone
33	448.4	19.3	519	18	AAZ02141	Human secreted pro
34	445	19.1	541	22	AAK92946	Human cDNA 3'-end
35	386	16.6	1480	21	AAZ14608	CDNA encoding a hu
36	386	16.6	1525	18	AAZ88974	Human extracellular
37	386	16.6	1531	20	AAZ60351	Human extracellular
38	386	16.6	1707	22	AAH13771	Human cDNA sequenc
39	386	16.6	1875	24	AAZ1402	Human cDNA sequenc
40	386	16.6	1875	24	ABK40256	CDNA encoding huma
41	386	16.6	1994	21	AAZ82239	Lung cancer associ
42	386	16.6	2018	21	AAZ39860	Human SI-5 RCM-11
43	382	16.4	1874	21	AAZ58583	Human PR226 prote
44	374.4	16.1	377	20	AAZ37686	Human EST sequence
45	352.4	15.1	1358	21	AAA14616	CDNA encoding the

ALIGNMENTS

RESULT 1	AAZ40027	AAZ40027 standard; DNA: 2328 BP.
XX	AAZ40027;	
AC	15-FEB-2000 (first entry)	
XX		
DT		
XX		
DE	Full length human A55 protein coding sequence.	
XX		
KW	A55 protein; human; smooth muscle proliferation; tissue generation;	
KW	vascular smooth muscle cell; arteriosclerosis; tissue repair; myoma;	
KW	percutaneous transluminal coronary angioplasty; blood coagulation; PTCA;	
KW	actin; inhibin; chemotaxis; thrombosis; cadherin; therapy;	
XX	tumour metastasis inhibitor; ss.	
OS	Homo sapiens.	
XX		
PN	WO9955864-A1.	
XX		
PD	04-NOV-1999.	
XX		
PF	28-APR-1999; 99WO-JP02284.	
XX		
PR	28-APR-1998; 98JP-0119731.	
XX		
PA	(ONOR) ONO PHARM CO LTD.	
XX		
PI	Honjo T, Tashiro K, Nakamura T;	
XX		
WPI	2000-038647/03.	


```
Db 90285 AAGTTTCAGTGTGCTTAATTTGTATACGGTTTAAATCTGCTGTTCATTTGAGTA 90344
OY 2133 TTTTAAAAAATATGTGTAGAAATTCCTTGAAAGGCTTCAGACATGCTATGTCTG 2192
    |||||||
Db 90345 TTTTAAAAAATATGTGTAGAAATTCCTTGAAAGGCTTCAGACATGCTATGTCTG 90404
    |||||||
OY 2193 TCTTCCAAACCCAGTCTCTCTCCATTTTAGCCCAAGTGTTCCTTGAGGACCCCTTAA 2252
    |||||||
Db 90405 TCTTCCAAACCCAGTCTCTCTCCATTTTAGCCCAAGTGTTCCTTGAGGACCCCTTAA 90464
    |||||||
OY 2253 TCTTCTTCTTCTTGTAGAAATTTTACCAATTTGATGGAATGCAAGAGTCTCCAAACTGAT 2312
    |||||||
Db 90465 TCTTCTTCTTCTTGTAGAAATTTTACCAATTTGATGGAATGCAAGAGTCTCCAAACTGAT 90524
    |||||||
OY 2313 TAAATATTGAGAGA 2328
    |||||||
Db 90525 TAAATATTGAGAGA 90540
    |||||||
```

Search completed: July 3, 2003, 15:34:27
Job time : 6074.42 secs


```

OY 1953 GGAGTAGCTAGTGTGCTTTTGGCTACACAGAGAGCGCTATGTAACCAACACAGCAGG 2012
    |||||||
Db 16137 GGAGTAGCTAGTGTGCTTTTGGCTACACAGAGAGCGCTATGTAACCAACACAGCAGG 16078
OY 2013 ATCGAAGGTTTTAGAGAAATGTTTCAAAAACCATGCTGTATTTTCAACCATTAAG 2072
    |||||||
Db 16072 ATCGAAGGTTTTAGAGAAATGTTTCAAAAACCATGCTGTATTTTCAACCATTAAG 16018
OY 2073 AAGTTTCAGTGTGCTTTAATTTGTATACGCTTAATTCGTCTGTTCATTTTGAGTA 2132
    |||||||
Db 16017 AAGTTTCAGTGTGCTTTAATTTGTATACGCTTAATTCGTCTGTTCATTTTGAGTA 15958
OY 2133 TTTTAAAAAATATGCTGAGAAATTCCTGGAAGGCTTCAGACATGCTATGTGTG 2192
    |||||||
Db 15957 TTTTAAAAAATATGCTGAGAAATTCCTGGAAGGCTTCAGACATGCTATGTGTG 15898
OY 2193 TCTTCCCAACCCAGTCTCTCTTCATTTTATGCCAGTGTTCCTTTGAGAGACCCCTTAA 2252
    |||||||
Db 15897 TCTTCCCAACCCAGTCTCTCTTCATTTTATGCCAGTGTTCCTTTGAGAGACCCCTTAA 15838
OY 2253 TCTTGTCTTCTTATGATTTTATCCCAATGGATTGGAATGACAGAGCTCCAAAGCTAT 2312
    |||||||
Db 15837 TCTTGTCTTCTTATGATTTTATCCCAATGGATTGGAATGACAGAGCTCCAAAGCTAT 15778
OY 2313 TAAATATTTGAAGAGA 2328
    |||||||
Db 15777 TAAATATTTGAAGAGA 15762

```

```

RESULT 15
LOCUS   CNS07EG5      175210 bp    DNA        linear   PRI 25-JUN-2001
DEFINITION   Human chromosome 14 DNA sequence BAC R-818K5 of library RPCL-11
ACCESSION   AL590328
VERSION      AL590328.2  GI:14571668
KEYWORDS     HTG: HTGS_ACTIVEFIN.
SOURCE       human.
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE
AUTHORS     Heilig,R., Petit,O.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
            Brothier,P., Catillo,L., Barbe,V., Pelletier,E., Artiguenave,F.,
            Levy,M., Eckenberg,R., Bruls,T., deBerardinis,V., Cruaud,C.,
            Gabyay,G., Saurin,W. and Weissenbach,J.
            Sequencing of the human chromosome 14
            Unpublished
            2 (bases 1 to 175210)
            Genoscope.
            Direct Submission
            Submitted (25-JUN-2001) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
            Web : www.genoscope.cns.fr)
            On Jun 26, 2001 this sequence version replaced gi:13509295.
            ----- Genome Center
            Center: Genoscope / Centre National de Sequencage
            Center code: GS
            Web site: http://www.genoscope.cns.fr/
            Contact: Sequef@genoscope.cns.fr

```

```

COMMENT
The following BAC sequence is oriented from the r7 to the sp6 end.
----- Summary Statistics
Assembly program: Phrap; version 2.0
Quality coverage: 7.46x in Q20 bases; sum-of-contigs

```

```

Overall quality chart :
Range      : bases
0          : 6
1 - 9     : 52
10 - 19   : 268
20 - 29   : 576

```

```

FEATURES
source
1..175210
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="14"
/clone="R-818K5"
/clone_id="RPCL-11"
BASE COUNT  46924 a 36746 c 38389 g 53145 t      6 others
ORIGIN
Query Match      41.9%  Score 974.4; DB 9; Length 175210;
Best Local Similarity 99.9%; Pred. No. 3.3e-261;
Matches 975; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1353 GCAAGCGGCCCATCATGATGCGACCTGCTGATGACAGCGCCCATCAAAAGGCGCCGGGA 1412
    |||||||
Db 89565 GCAAGCGGCCCATCATGATGCGACCTGCTGATGACAGCGCCCATCAAAAGGCGCCGGGA 89624
OY 1413 AATTCAGCTGACTTGGAAATGATCATCTGTCAACACTGTCTATCAACTTCAGAGCGACTC 1472
    |||||||
Db 89625 AATTCAGCTGACTTGGAAATGATCATCTGTCAACACTGTCTATCAACTTCAGAGCGACTC 89684
OY 1473 CGTGATCGAGTGGGATATATGTGTGTCGACAGTACCATTTCTGAGCTGCGGCTGGAGCCT 1532
    |||||||
Db 89685 CGTGATCGAGTGGGATATATGTGTGTCGACAGTACCATTTCTGAGCTGCGGCTGGAGCCT 89744
OY 1533 CCGAGCGTGCCTTCATTTGGCAGCAAGGAGAGAGAGAGAAATTAACAGAGAGAT 1592
    |||||||
Db 89745 CCGAGCGTGCCTTCATTTGGCAGCAAGGAGAGAGAGAGAAATTAACAGAGAGAT 89804
OY 1593 GAGAGCGACAGAGCGTTTATGAGCATTTCTGCTGAAAGCTTTCCGCAAGATGACGCCGA 1652
    |||||||
Db 89805 GAGAGCGACAGAGCGTTTATGAGCATTTCTGCTGAAAGCTTTCCGCAAGATGACGCCGA 89864
OY 1653 CTTCCTGACCTGCTACCTGCTACTATTTGACAGCTGTACCTGCTGACAGACTTGGCACCCCA 1712
    |||||||
Db 89865 CTTCCTGACCTGCTACCTGCTACTATTTGACAGCTGTACCTGCTGACAGACTTGGCACCCCA 89924
OY 1713 GTTCTATGATACAGTTATCAAAAAGTATATCATTTGCTCCCTGATAGAAATTTGTTGG 1772
    |||||||
Db 89925 GTTCTATGATACAGTTATCAAAAAGTATATCATTTGCTCCCTGATAGAAATTTGTTGG 89984
OY 1773 TGAATTTTCAAGGCTTCAGTTATTTCCACTATTTTCAAAAGAAATGATTAAGTTTGC 1832
    |||||||
Db 89985 TGAATTTTCAAGGCTTCAGTTATTTCCACTATTTTCAAAAGAAATGATTAAGTTTGC 90044
OY 1833 GGGGGTCTGAGTCTATGTTTCAAAAGCTGTGAACAGCTGTCTGCTCTTCAACCTCTTC 1892
    |||||||
Db 90045 GGGGGTCTGAGTCTATGTTTCAAAAGCTGTGAACAGCTGTCTGCTCTTCAACCTCTTC 90104
OY 1893 CACTCTCTCTCTACCTGTGTATACGCTTTGCAAAAGACCCGGAGGTGCGGGGAACCTCG 1952
    |||||||
Db 90105 CACTCTCTCTCTACCTGTGTATGCTTTGCAAAAGACCCGGAGGTGCGGGGAACCTCG 90164
OY 1953 GGAGTAGCTAGTGTGCTTTTGGCTACACAGAGAGCGCTATGTAACCAACACAGCAGG 2012
    |||||||
Db 90165 GGAGTAGCTAGTGTGCTTTTGGCTACACAGAGAGCGCTATGTAACCAACACAGCAGG 90224
OY 2013 ATCGAAGGTTTTAGAGAAATGTTTCAAAAACCATGCTGTATTTTCAACCATTAAG 2072
    |||||||
Db 90225 ATCGAAGGTTTTAGAGAAATGTTTCAAAAACCATGCTGTATTTTCAACCATTAAG 90284
OY 2073 AAGTTTCAGTGTGCTTTAATTTGTATACGCTTAATTCGTCTGTTCATTTTGAGTA 2132
    |||||||

```

QY 1672 ACTATTGACAGCTGTCACCCCTGACAGACTTGGCACCCTGCTTATGATACAGTTAT 1731
 Db 1653 ACTATCGCAGA-CTGTCACCTCTGACAGACACCAATGCCCTGCTTATGATGACAGTTAT 1711
 QY 1732 CAAAAGATATATATGCTGCTCCCTGATAGAGATGTTGGTGAATTTTCAAGCCCTTCA 1791
 Db 1712 CCAAAAGATATCTCTG-ACCCTGATGAGGTTCAAGTACCTCCAGGCCCTTCC 1770
 QY 1792 GTTATTTCACATATTT-----CAAGAAAATAGATTAGTTTCCGGGGCTGAGT 1844
 Db 1771 ATTATTCTATATTTTATTTTAAAGAAAATAGATTAGTTTCTGGGGTGAAGT 1830
 QY 1845 CTATGCTCAAGAGCTGTAAGCCTGCTGCTCACTGCTTCACTCCCTCCCTGCT 1904
 Db 1831 CTGCTTCAAGAC-TCAGCAGCTTGTCTTCTGCTTCCCTCCCTCTTC-CTATATCT 1887
 QY 1905 CACTGTGTACTGCTTGGAAA--GACCCGGAGCTGGCGGGAAACCTGGAGTAGTA 1962
 Db 1888 TGCTCATGTGCTGCTTGGAAAAGTCTCATGAGCTCATGGAATGCTGGGTATAGTA 1947
 QY 1963 GTTTCCTTTTCCGTCACAGAGAGGCTATGTAACAAACACAGAGATGCAAGGT 2022
 Db 1948 TTTTCCTTCTGCTGCTGCTGAGAAAGCTATGGGAAACGCTCACAGCAGATC-AAAGGT 2006
 QY 2023 TTTTGAAGATGTGTTCAAAAACCATGCTGATTTTCAACCATTAAGAAGTTTGT 2082
 Db 2007 TTTTAAAGATGATTTTAAATTCAGCTGCTGATTTTTCAGCATTAAGAAAGTTTGT 2066
 QY 2083 TGTCTTAAATTTGATTAAGGTTTAACTGCTGCTGCTGCTGCTGCTGCTTAA 2142
 Db 2067 GTTCTTAAATTTGATGAGTGTAACTTCTGCTGCTGCTGCTGCTGCTTAA 2119
 QY 2143 ATATGTCGAGATTCCTCGAAAGGCTTCACACATCTATGCTGCTGCTGCTGCT 2202
 Db 2120 -AAAGTATAGATTCCTCGAAAGGCTTCACATCTATGCTGCTGCTGCTGCT 2178
 QY 2203 CCAAGCTCTCTCTCATTTTACCCAGTGTCTTCTGAGAGCCCTTAACTGCTTTC 2262
 Db 2179 CTCACCTTCTCTGCTGCTTACGCCA--GTTTTATGAGAACCCCTTAACTGCTTTC 2236
 QY 2283 TTTAGATTTTACCAGATTTGAAATGCAAGGCTTCACAACTGATTAATATTTG 2322
 Db 2237 TTAAGAGTTTATACCACTGGGTTGGAAGACAGAGTATCCGACTGATTAATATTTG 2296
 QY 2323 AAGAGA 2328
 Db 2297 AAGAAA 2302
 RESULT 14
 AC007057/c 171901 bp DNA linear PRI 25-APR-1999
 DEFINITION Homo sapiens chromosome 14 clone containing genes for UP50 and
 AC007057
 AC007057
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 171901)
 Rowen, L., Madan, A., Qin, S., Abbasi, N., Dors, M., Dickhoff, R.,
 James, R., Loretz, C., Lasky, S., Madan, A., Prescott, S., Ratcliffe, A.,
 Shafer, T., and Hood, L.
 Sequencing of human chromosome 14 gene for UP50 and partial gene
 for Trip230, thyroid hormone receptor co-activator
 Unpublished
 2 (bases 1 to 171901)
 Rowen, L., Madan, A., Qin, S., Abbasi, N., Dors, M., Dickhoff, R.,
 Hicks, P., James, R., Loretz, C., Lasky, S., Madan, A., Ratcliffe, A.,
 Shafer, T., and Hood, L.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 FEATURES
 source
 Direct Submission
 Submitted (11-MAR-1999) Multimegabase Sequencing Center, University
 of Washington, PO Box 357730, Seattle, WA 98195, USA
 3 (bases 1 to 171901)
 Rowen, L., Madan, A., Qin, S., Abbasi, N., Dors, M., Dickhoff, R.,
 James, R., Loretz, C., Lasky, S., Madan, A., Prescott, S., Ratcliffe, A.,
 Shafer, T., and Hood, L.
 Direct Submission
 Submitted (25-APR-1999) Multimegabase Sequencing Center, University
 of Washington, PO Box 357730, Seattle, WA 98195, USA
 On Apr 25, 1999 this sequence version replaced gl:4558634.
 Location/Qualifiers
 1. 171901
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="14"
 /map="14q31"
 /clone="BAC 353N19"
 /clone_lib="RPC111"
 complement(16075..93737)
 /gene="UP50"
 /note="The cDNA for this gene is found in AF093118."
 complement(115917..171768)
 /gene="Trip230"
 /note="The cDNA for this gene is found in AF007321. The
 5' end of the gene is missing from this clone."
 BASE COUNT 51278 a 38116 c 35671 g 46836 t
 ORIGIN
 Query Match 41.9%; Score 974.4; DB 9; Length 171901;
 Best Local Similarity 99.9%; Pred. No. 3.3e-261;
 Matches 975; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1353 GCAAGGCGCCCATATGTCACACCTGCTGATGACACGCCATCAAGGCGCCCGGA 1412
 Db 16737 GCAAAGGCGCCCATATGTCACACCTGCTGATGACACGCCATCAAGGCGCCCGGA 16678
 QY 1413 AATCCAGTGGACTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1472
 Db 16677 AATCCAGTGGACTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 16618
 QY 1473 CGTATCCAGTGGACTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1532
 Db 16617 CGTATCCAGTGGACTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 16558
 QY 1533 CCGAGCCTCCTCTCATTTGGCCACCAAGGACAGAGAAAGAGAAATACAGAGAT 1592
 Db 16557 CCGAGCCTCCTCTCATTTGGCCACCAAGGACAGAGAAAGAGAAATACAGAGAT 16498
 QY 1593 GAGAGGACACAGAGCTTAGGCAATTCCTGTAAGCTTCCCGGAAGATCAGCCCGGA 1652
 Db 16497 GAGAGGACACAGAGCTTAGGCAATTCCTGTAAGCTTCCCGGAAGATCAGCCCGGA 16438
 QY 1653 CTTCCTGACTCTACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1712
 Db 16437 CTTCCTGACTCTACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 16378
 QY 1713 GTTCTATGATACAGTATTAACAAAAGATTAATATGCTGCTGCTGCTGCTGCTGCT 1772
 Db 16377 GTTCTATGATACAGTATTAACAAAAGATTAATATGCTGCTGCTGCTGCTGCTGCT 16318
 QY 1773 TGAATTTTCAAGGCTTCAAGTATTAATTTTCAAGTATTAATTTTCAAGTATTAAT 1832
 Db 16317 TGAATTTTCAAGGCTTCAAGTATTAATTTTCAAGTATTAATTTTCAAGTATTAAT 16258
 QY 1833 GGGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1892
 Db 16257 GGGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 16198
 QY 1893 CACTCCTTCTCTCACTGCTTACTGCTTTCGAAAGACCCGGAGACTGGCGGGAACCTG 1952
 Db 16197 CACTCCTTCTCTCACTGCTTACTGCTTTCGAAAGACCCGGAGACTGGCGGGAACCTG 16138

QY	352	CAAAATGGGGGGATTTATATGACATTCCTCCCGSAGCAAACCCCTGTGATGTAGAGGGCCCTTACTGC	411
Db	533	CAGATATGGGGGATATCTGTGACATCCCTCGAACCAACCCAGTGTATGAGAGGGCCCTTACTGC	592
QY	412	AACCCCTACTCGACCCCTACTCAGGGTCCGTACCCAGCAGCTGCCCCACACTCTCAGCT	471
Db	593	AATCCCTACTCTCATCTTACTCAGGCCATACCCAGCAGCCGACACCCAGCTGCACGT	652
QY	472	CCAAATATCCCAACGATCTCCAGGCTCTTATATATGCTCCGCTTTGGATACCATGTGAA	531
Db	653	TCCAACTATCCCAACGATCTCCAGGCTCTTGTCTGTGCTTGGGTATCAGATGTATGAA	712
QY	532	AGCAACCAATGTGTGATGTGACGAGTGTGCAACGATATCCACAGCGCAACCCACAC	591
Db	713	GGCAACCAATGTGTGATGTGACGAGTGTGCAACGATATCCACAGCGCAACCCACCTTAC	772
QY	592	CAGATCTGCATCAATATCTGAAGCGGGATCACCTGCTCCTGCAACGACGAGATTTGGCT	651
Db	773	CAGATCTGTATCAACCGGAGAGGGATCACCTGCTCCTGCAACGATTTGGCT	832
QY	652	CTGGAAGGCGCTGCTTACATTTGATGTAATGCCATATGTTACTGCGACGACGCTCTGT	711
Db	833	CTGGAAGGCGCTTACATTTGATGTAATGCCATATGTTACTGCGACGACGCTCTGT	892
QY	712	CGGAATGTTCTTGATTCCTATTCTTGTATACATCAACCCCTGTTTTACCTCAATGAGAT	771
Db	893	CGGAATGTTCTTGATTCCTATTCTTGTATACATCAACCCCTGTTTTACCTCAATGAGAT	952
QY	772	GGAAAGTCTTGCCAAAGATGTGAAGCGAGTGTGCACCGAGAACCCCTGGGTGAACCTGC	831
Db	953	GGAAAGTCTTGCCAAAGATGTGAAGCGAGTGTGAACCTGGAGAACCCCTGGGTGAACCTGC	1011
QY	832	GTCAACACCTATCGCTCTTTCATCTGCGCTGTGACCCAGATATGAATGAATGGAGAAAT	891
Db	1013	GTCAACACCTATCGCTCTTTCATCTGCGCTGTGACCCAGATATGAATGAATGGAGAAAT	1072
QY	892	GGCCTTCAATGCAAGTATATGAGAGTGAAGAGCTTCTGTGATTTCTTGCACATATGAG	951
Db	1073	GGCCTTCAATGCAAGTATATGAGAGTGAAGAGCTTCTGTGATTTCTTGCACATATGAG	1132
QY	952	TGTGTGAACACGCGCGGACATATCTTGTCTCTGCTCCCTCGACGCTATCTCTGTGAT	1011
Db	1133	TGTGTGAACACGCGCGGCTCATCTTGTCTCTGCTCCCTCGACGCTATCTCTGTGAT	1192
QY	1012	GACAACCGAATCTGCCAAGACATCAACAATGTGACACAGAGAACACACAGCTGCACCTG	1071
Db	1193	GATAACCGAATCTGCCAAGATCAACAATGTGACACCGAGAACACACATGCACTGCC	1252
QY	1072	CAGCAGAGTGTCTACAAATTTACAAGGGGGCTTCAAAATGCATCGAACCCCATCGCGTGA	1131
Db	1253	CTGCAGATTTGTCAAAATCTGCAGAGGGGGCTTCAAAATGTATTCGACCCCATGTCTGCAG	1312
QY	1132	GAGCCTTATCTGAGGATCATGTATACCGCTGTATGTGTCTCTGTGAGAACCTTGGCTGC	1191
Db	1313	GAGCCTTATCTGAGGATTTGGGGATAAACCGCTGTATGTCTCTCTGTGAGAAATAGGGCTGC	1372
QY	1192	AGAGACAGCCCTTATACCTTTGTACCGGGGAATGAGCGTGTGTGACGACGCTCCGT	1251
Db	1373	AGGACACAGCCATTCACCATTTTGTTCGGGACATGAGTGTGTGTATGAGCGCTGTGT	1432
QY	1252	CCCGCTGACATCTTCCAATTCGAAGCCAGACCGCGCTACCTGTGGGGCTTATTAATTTTC	1311
Db	1433	CCTGTGTACATCTTCCAGATGCAGAACGACCGGATACCTGTGGCGCTTATTAATTTTC	1492
QY	1312	CAGATCAAAATCTGGGAATGAGGGCAGAGAAATTTTACATGCGGCAACGGGCCCATCATGT	1371
Db	1493	CAGATCAAAATCTGGGAACGAGGGTCCAGAGTTTCAATATGCGGCAACAGGGCCCTATCATGT	1552
QY	1372	GCCACCTGTGTGAATGACAGCGCCCATCAAAAGGGGCCGGGAAATTCACGCTGGGAATTTGAA	1431
Db	1553	GCCACCTGTGTGAATGACAGCGCCCATCAAAAGGGCCCTGGGCAATCCACGCTGGGAATTTGAG	1612
QY	1432	ATGATCACTGTCAACACTGTATCAACTTCAAGAGGACGCTCCGTGATCCGACTCCGATA	1491

Accession	Source	Position	Sequence	Length
Db	AF137350	1613	ATATATCAGCGTCAACACTGTCAATCACTTACAGAGGAGCTCCGTGATCCGACTGGGAGATA	1672
Qy	AF137350	1492	TATGTGCGAGACACCATTCTGAGAGCGTCGGGGCTGGAGGCTCCGACAGCTGCTCTCATTTG	1551
Db	AF137350	1673	TACGTGTCCCAAGTATCCGTTCTGAGGCTGTGGGTTAAGGCTCTGACACAGCTCTTTTACCA	1733
Qy	AF137350	1552	GCACCAAGGAGACAGAGAGAGAGAAATACAGAGAGATGAGAGCGACACAGAGCTTA	1611
Db	AF137350	1733	GCACCGAGGAGACAGAGAGAGAGAA - ACACAGAGAAATAGAGCGGAGAGAGAGAGATTTG	1790
Qy	AF137350	1612	GGATTTCTCTGCTGAACGTTTCCCGAAGATCAGCCCGACTTCCGTGACTCTACCTGT	1677
Db	AF137350	1791	CACCTTTCCTGCTGAACNTCTCCCTGGGACATCAGCTAG - CATCTGACCCCTACCTGT	1845
Qy	AF137350	1672	ACTATGACAGACCTGACCTTGCAGAGACTTCCACCCCGACTTCCGTATGATACAGTAT	1733
Db	AF137350	1850	ACATGCGAGA - CTGCACTCTGAGAGACACATGCCCCCACTTCTATGATGACAGTAT	1906
Qy	AF137350	1732	CAAAAAGTATATCATGTCTCCCTGTATAGAAATGTTGTAATTTTCAAGGCTTCA	1791
Db	AF137350	1909	CCAAAAGTATCTCTTG - ACCCTGATGATGAGTGAAGAGATCTCCAAAGGCTTCC	1967
Qy	AF137350	1792	GTTATTTTCCACTATTTT - CAAAGAAAATAGATTGTTGTTGGGGGGCTAG	1845
Db	AF137350	1968	ATTATTTTCTATTTTTTTTTTTTTTAAAAAGAAAATACATTATGTTTCTGGGGTAGAG	2022
Qy	AF137350	1844	TCTATGTTCAAGAGCTGTGACAGCTTGTCTACTTCTTCACTCTTCCACTCTTCTC	1903
Db	AF137350	2028	TCTGTTTCAAAAGAC - TGAGAGAGCTTGTCTTGGCTCTGCCCCCTTTC - CTATATCTC	2084
Qy	AF137350	1904	TCACTGTGTACTGCTTTGCAAA - GACCCGGGAGCTGGGGGAGACCTCGGGAGTACT	1961
Db	AF137350	2085	TTCGTGCAATGCTGCTTTGCAAAATCTCTCATGAGCTATGAGAAATCTGGGATAGCT	2144
Qy	AF137350	1962	ACTTTGCTTTTGGCTACACAGAGAGGCTATGTAACAAACACACAGAGATCGAAGG	2021
Db	AF137350	2145	ATTTTGTCTTTCATGCTGCTGAGAGGCTATGAGAACACGTCACAGAGATC - AAAGG	2203
Qy	AF137350	2022	TTTTTAGAGAAATGTTTCAAAACCATGCTGATTTTTCACATTAATAAGATTCTAG	2081
Db	AF137350	2204	TTTTTAAGAGCTCTATTTAAATACAGCTGTGATTTTTCAGCATTAAGAAAGTTTAACT	2263
Qy	AF137350	2082	TTGTCTTAAATTTGCTATACAGGTTTAAATCTGCTGCTATTTTGAATTTTAAAA	2141
Db	AF137350	2264	TGTTCTTAAATTTGATGAGGTTTAACTTTCTGCTTCTATTTTGAATTTT -	2317
Qy	AF137350	2142	AATATGCTGATGAATCTCTCGAAAGGCTTACAGACATCTATGTTCTGTCTCCAA	2201
Db	AF137350	2318	--AAAGTATGAATTTCTTCCAGAGGCTTACAGATATGTTATGTTCACTTTTCAA	2375
Qy	AF137350	2202	ACCAGTCTCTCTCATTTTACGCAATGTTTCTTGGAGAGACCCCTTAATCTTGGTT	2261
Db	AF137350	2376	CCTGAGCTTCTCTGCGTCTTACCCA - GTTTTATGAGAACCCCTTTAATCATGCTTT	2433
Qy	AF137350	2262	CTTTAGATTTTTCACCAATGATGATGATGACAGAGGCTCCAACTGATTAATATTT	2321
Db	AF137350	2434	CTTAAGATTTTTCACCAACTGAGGTTGAGACAGAGAGATCCGCACTGATTAATATTT	2493
Qy	AF137350	2332	GAGA 2326	
Db	AF137350	2494	GAGA 2498	

1313 AGATCAATCTGGGAATGAGGCGAGAAATTTTACATCGGCAACGGGCCCATCTAGT 1372
 1464 AGATCAATCTGGCAACGAGGCGTCAATATATGCGGCAACAGGCGCTATCACTG 1523
 1373 CCACCCCTGGTATGACAGCGCCCATCAAAAGGCGCGGAAATCACTGAGCTTGAAA 1432
 1524 CCACCCCTGGTATGACAGCGCCCATCAAAAGGCGCGGAAATCACTGAGCTTGAAA 1583
 1433 TGATCACTGCAACAGTCACTCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1492
 1584 TGATCACTGCAACAGTCACTCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1643
 1493 ATGTGCGCAGTACCCATTTGAGAGCTCGGAGCTGAGAGCTCGAGCTCGCTCATTTG 1552
 1644 ATGTGCGCAGTACCCATTTGAGAGCTCGGAGCTGAGAGCTCGAGCTCGCTCATTTG 1703
 1553 CACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1612
 1704 CACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1761
 1613 GCATTTCTGCTGAGAGCTTTCCCGGAGAGTCAAGCCCGAGCTTCTGAGCTCACTGTA 1672
 1762 ACCTTTCTGCTGAGAGTCTTCTGAGGAGGAGTCAAGCCCGAGCTTCTGAGCTCACTGTA 1820
 1673 CTATTCAGAGACCTGTCACCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1732
 1821 CTATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1879
 1733 AAAAAATATATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1792
 1880 CAAAAATATATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1938
 1793 TTTATTTTCCATATTTT---CAAGAAATATATATATGCTGCTGCTGCTGCTGCTGCTGCT 1846
 1939 TTTATTTTCCATATTTTAAAGAAATATATATATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1998
 1847 -ATGTCAAGAGCTGTCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1905
 1999 GAAAGTTCAAAAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2057
 1906 ACTGCTTACTGCTTGGCAAA--GACCCGGAGAGTGGCGGGGAAACCTGGGAGTACTAG 1963
 2058 GCTGATGCTGCTTGGCAAAAGTCCCTCATGAGGCTGCTGCGGAAATGCTGGGAAATAGCTAG 2117
 1964 TTTGCTTTTGGGTACAGAGAGAGGCTATGTAACAAACACAGCAGATCGAAGGCT 2023
 2118 TTTGCTTTTGGGTACAGAGAGGCTATGTAACAAACACAGCAGATCGAAGGCT 2176
 2024 TTTAGAGATGCTTTCAAACAGCAGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2083
 2177 TTTATAGATGCTATTTTAAATCAATCTGTAATTTTACACATTAAGAAATTTTACTG 2236
 2084 GTCCTTAATTTGTAATACGCTTAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2143
 2237 TCTTAAATTTGTAATGAGTGTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACT 2293
 2144 TATGCTGAGATTTCTTGAAGAGGCTTCAACAGCAGATGTAATGCTGCTGCTGCTGCTGCT 2203
 2294 ---GTGGTATGATTTCTTGAAGAGGCTTCAACAGCAGATGTAATGCTGCTGCTGCTGCT 2349
 2204 CCAGTCTCTCTCAATTTTGAAGAGGCTTCTTGAAGAGGCTTCAACAGCAGATGTAATGCTGCT 2263
 2350 TCATCTCTCTCTCAATTTTGAAGAGGCTTCTTGAAGAGGCTTCAACAGCAGATGTAATGCTGCT 2406
 2264 TTAGAATTTTATCCCAATTTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 2323
 2407 TAAAGATTTTATCCCAATTTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 2466
 2324 AGAGA 2328
 2467 AGAAA 2471

RESULT 12
 AF112153
 LOCUS
 DEFINITION
 Rattus norvegicus developmental arteries and neural crest EGF-like protein mRNA, complete cds.
 ACCESSION
 AF112153
 VERSION
 AF112153.1
 KEYWORDS
 ORGANISM
 Rattus norvegicus.
 Rattus norvegicus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 REFERENCE
 1 (bases 1 to 2499)
 Nakamura, T., Ruiz-Lozano, P., Lindner, V., Yabe, D., Taniwaki, M., Furukawa, Y., Koduke, K., Tashiro, K., Lu, Z., Andon, N.L., Schaub, R., Matsumori, A., Sasayama, S., Chien, K.R. and Honjo, T.
 DANCE, a novel secreted RGD protein expressed in developing, atherosclerotic, and balloon-injured arteries
 J. Biol. Chem. 274 (32), 22476-22483 (1999)
 JOURNAL
 MEDLINE
 PUBMED
 10428823
 2 (bases 1 to 2499)
 Nakamura, T., Yabe, D., Tashiro, K. and Honjo, T.
 Direct Submission
 Submitted (09-DEC-1998) Medicine, University of California San Diego, 9500 Gilman drive, La Jolla, CA 92093-0613, USA
 JOURNAL
 FEATUES
 source
 Location/Qualifiers
 1. 2499
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 350. 1696
 /note="DANCE: contains 6 cbeGF domains and 1 RGD motif: secreted protein; expressed in developmental aorta, neural crest cells, balloon injured vessels, atherosclerotic lesions"
 /product="developmental arteries and neural crest EGF-like protein"
 /protein_id="AA041769.1"
 /db_xref="GI:5305675"
 /translation="MPELRILVTITLALMPLRFGMAOQCTNGFDLDRGTGCLDID ECHTIEACGDMACVNGNGYLCPITNPVRYSPYSTYSYPAAPVPAPAS NPTILSRPLVCRFGYOMDEGNOCVDVDEACDSHOCTOICINEGGTCGTGYW LLEGCLDIDECRYGCOOLCANPAGSCNCFNTINDGRSDOYNECTEUPCV QTVVNTGSRFCRDPETELBEDGIRHSDMECSFSEFLCQHECVNPGSTYSCPPG YVLEENRSCQDINECHRNHHTCPIDTCYNLDGFKCIDVCEPEYLLIDNRCMC PAENTGCRDPTILFERDMDVSGRSVPADIFQOATRYRGAYYIFIKSGNEBREF YMRGTGPIASITLVMTPIKGPRIQDLDEMIVTIVINFRGSSVIRLRIYSQYFP"
 BASE COUNT
 584 a 669 c 595 g 649 t 2 others
 ORIGIN
 Query Match 63.48; Score 1476.6; DB 10; Length 2499;
 Best Local Similarity 82.7%; Pred. No. 0;
 Matches 1839; Conservative 0; Mismatches 356; Indels 30; Gaps 12;
 Db 112 GGTCCCGCGCGAGAGCTTTCTGCGCTTGCAGATCTCTGCGGCGGTGGAGCATG 171
 294 GGTCCCGCGAGAGAGCTTCTCTCAATATACCTCTCTCC-GCGGATCTGGATATG 352
 172 CCAGGATTAAGAGATCTACTGTTACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 231
 353 CCAGGATTAAGAGATCTACTGTTACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 412
 232 AATGACAGGACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGT 291
 413 AATGACAGGACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGT 472
 292 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 351
 473 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 532

```

Db      2090 TCATCCCTTCTCGCATCTTACCCAGTTTAC---GAAGACCCCTTATCATCGCTTCT 2146
Oy      2264 TTGAATTTTACCAATGGATTGGAAATGCAAGAGTGTCCAACTGATTAATATTTGA 2323
Db      2147 TAAGAGTTTACCACTGCGTTGGAAAGACAGAGTATCCAGCTGATTAATATTTGA 2206
Oy      2324 AGAGA 2328
Db      2207 AAAAA 2211

RESULT 11
AF112151
LOCUS    AF112151
DEFINITION Mus musculus developmental arteries and neural crest EGF-like
ACCESSION AF112151
VERSION   AF112151.1 GI:5305670
KEYWORDS
SOURCE    Mus musculus.
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2478)
AUTHORS  Nakamura,T., Kobuke,K., Tashiro,K., Lu,Z., Andon,N.L., Schaub,R.,
          Matsumori,A., Sasayama,S., Chien,K.R. and Honjo,T.
          DANCE, a novel secreted RGD protein expressed in developing,
          atherosclerotic, and balloon-injured arteries
          J. Biol. Chem. 274 (32), 22476-22483 (1999)
JOURNAL  2 (bases 1 to 2478)
MEDLINE  10428823
PUBMED   99357779
REFERENCE 2 (bases 1 to 2478)
AUTHORS  Nakamura,T., Yabe,D., Tashiro,K. and Honjo,T.
          Direct Submission
          Submitted (09-DEC-1998) Medicine, University of California San
          Diego, 9500 Gilman Drive, La Jolla, CA 92093-0613, USA
FEATURES
source    location/Qualifiers
          1..2478
          /organism="Mus musculus"
          /db_xref="taxon:10090"
          /note="DANCE: contains 6 cbegf domains and 1 RGD motif;
          secreted protein; expressed in developmental aorta, neural
          crest cells, balloon injured vessels, atherosclerotic
          lesions"
          /codon_start=1
          /product="developmental arteries and neural crest EGF-like
          protein"
          /protein_id="AA041767.1"
          /db_xref="GI:5305671"
          /translation="MPELKRLLVITLALMPLPHENAOQCTNEDLDROSQCCLDID
          ECHTIPACKGMVNVNNGGYLCIPRNPYRGPYSYPSTYSGPVPAAPVPAS
          NYPTISRLVREGYOMDEGNOCVDECATDSHCNFTQICINTEGVTSCITGYW
          LLEGQCLDIDECRGYCOOLCANPAGSYSCNPGFTINDGRCQVDECTENPCV
          QTVNTYSGFICRCDPGELEEDGICHSDEGSEFELCGHECVNPGSYSCSPG
          YVLLDNRSQDQIDNECHRNHTCSLOTGLOGFKIDPISCEPELILGENCMC
          PAHTSCRDQPTFLYRDMVYSGRSPADIFQMATTRTGTATTFEIKSNEGREF
          YMRGTGISLTVTRPIKGRDITQLEMTVNTVINFRSSVIRLRIVSYQYP"

BASE COUNT 604 a 656 c 575 g 643 t

Query Match 63.7% Score 1481.8; DB 10; Length 2478;
Best Local Similarity 82.8% Pred. No. 0;
Matches 1842; Conservative 0; Mismatches 357; Indels 26; Gaps 12;

```

```

Db      324 CAGGATTAAGAGATACACGTCTTACCATCTTGGACAGCTGTGGCTCCACATCCCGGA 383
Oy      233 ATGCACAGGACAGTGCACGATGGCTTTGACCTTGATGCCAGTGCAGACAGTGTAG 292
Db      384 ATGCACAGGACAGTGCACGATGGCTTTGACCTTGATGCCAGTGCAGACAGTGTAG 443
Oy      293 ATATGTGATGATGCGGACCAATCCCGGCTGCGGAGAGACATGATGTGTAAAC 352
Db      444 ATATGTGATGATGCGGACCAATCCCGGCTGCGGAGACATGATGTGTCAACC 503
Oy      353 AAATGGCGGATATTTATGATTCGCCGACAAACCCGTGTATGAGGGCCCTACGA 412
Db      504 AGATGGCGGATATTTGATGATTCGCCGACAAACCCGTGTATGAGGGCCCTACGA 563
Oy      413 ACCCTTACTGACCCCTTACTCAGGTCCGTGCCAGACAGCTGCCACACTCTACGTC 472
Db      564 ATCCCTACTTACTTCTTACTCAGGCCCATCCAGCAGCGGCCCCACAGTACCAAGT 623
Oy      473 CAACATGCCAGATCTCCAGGCTCTTATATGCGGCTTGGATACCAAGATGATGAA 532
Db      624 CCAATACCCACAGATTTCAAGGCTCTTGTCTGCGCTTGGATATGATGATGANG 683
Oy      533 GCAACCAATGTGTGATGTGAGACAGTGTGCAACAGATTCACACAGTGCACACCCACC 592
Db      684 GCAACCAAGTGTGATGTGAGACAGTGTGCAACAGATTCACACAGTGCACACCCACC 743
Oy      593 AGATGTGCATCATACGTAAGGCGGCTACACTGCTCTCTGACCGACGAGATTTGGCTTC 652
Db      744 AGATGTGCATCATACGTAAGGCGGCTACACTGCTCTCTGACCGATGGTACTGTGCTC 803
Oy      653 TGAAGGCGAGTGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATG 712
Db      804 TGAAGGCGAGTGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATG 863
Oy      713 CGAATGCTCTGATCTTATCTTGTACATGCAACCCGTGTTTACCTCAATGAGATG 772
Db      864 CAATGTCTCAGATCTTATCTTGTACATGCAACCCGTGTTTACCTCAACGATG 923
Oy      773 GAAGGCTTCCCAAGATGTGAGAGTGTGACGAGTGTGACGAGTGTGACGAGTGTG 832
Db      924 GAAGGCTTCCCAAGATGTGAGAGTGTGACGAGTGTGACGAGTGTGACGAGTGTG 983
Oy      833 TCAACACCTTACGCTCTTATCTTGTGACGAGTGTGACGAGTGTGACGAGTGTG 892
Db      984 TCAACACCTTACGCTCTTATCTTGTGACGAGTGTGACGAGTGTGACGAGTGTG 1043
Oy      893 GCGTTCTTCAAGTATGATGAGAGTGTGACGAGTGTGACGAGTGTGACGAGTGTG 952
Db      1044 GCATTTACTGACAGTATGATGAGAGTGTGACGAGTGTGACGAGTGTGACGAGT 1103
Oy      953 GTGTGAACGACGCGGACATATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1012
Db      1104 GTGTGAACGACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1163
Oy      1013 ACAACGACGCTGCAAGACATCAACGAATGTGACAGACAGGACACAGTGCACACTGC 1072
Db      1164 ATACGACGACGCTGCAAGACATCAACGAATGTGACAGACAGGACACAGTGCACACTGC 1223
Oy      1073 AGCAGACGCTGCAATTTTACAGAGGAGCTTCAATATGATGACACCCCATCCGTGAGG 1132
Db      1224 TGCAGACTTCTCAATCTCAAGAGGAGCTTCAATATGATGATGATGATGATGATG 1283
Oy      1133 AGCCTTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1192
Db      1284 AGCCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1343
Oy      1193 GAGACGACGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1252
Db      1344 GAGACGACGATTCATCATCTCTGATGAGGACATGATGATGATGATGATGATGATG 1403
Oy      1253 CCGCTGACATCTTCAATATGACGACGACGACGACGACGACGACGACGACGACGAC 1312
Db      1404 CTGCTGACATCTTCAATATGACGACGACGACGACGACGACGACGACGACGACGAC 1463

```


Dp	525	CACGATCTCAGGCGCTTATATGCGGGTTTGATACCGAGATGATGAAAGCAACCAT	584
Qy	543	TGTGGATGTGGACGAGTGTGCAACAGATTTCCACCAATGCAACCCACCAGATTCGAT	602
Dp	585	TGTGGATGTGGACGAGTGTGCAACAGATTTCCACCAATGCAACCCACCAGATTCGAT	644
Qy	603	CAATACTGNAAGCGGGTACACCTGCTCTGCAACGACGAAATTTGGCTTTGGAAGGCCA	662
Dp	645	CAATACTGNAAGCGGGTACACCTGCTCTGCAACGACGAAATTTGGCTTTGGAAGGCCA	704
Qy	663	GTGCTTAGACATTTGATGAATGTGCTATGTATTACTGCCAGCAGCTGTGTGGATGTTCC	722
Dp	705	GTGCTTAGACATTTGATGAATGTGCTATGTATTACTGCCAGCAGCTGTGTGGATGTTCC	764
Qy	723	TGGATCCATTTCTTACATGCAACCGGGTTTTTCCCTCAATGAGATGGAAGTCTTG	782
Dp	765	TGGATCCATTTCTTACATGCAACCGGGTTTTTCCCTCAATGAGATGGAAGTCTTG	824
Qy	783	CCAAAGATGTGAACGAGTGTGCCACACGAGAACCCCTGCGTGCMAACCTCGTCACACCTA	842
Dp	825	CCAAAGATGTGAACGAGTGTGCCACACGAGAACCCCTGCGTGCMAACCTCGTCACACCTA	884
Qy	843	CGGCTCTTTCATCTGCGGCTGTGACCCAGATATGAACTTGAAGAAATGGCGTTCAATG	902
Dp	885	CGGCTCTTTCATCTGCGGCTGTGACCCAGATATGAACTTGAAGAAATGGCGTTCAATG	944
Qy	903	CAGTGATATGAGACGAGTGCAGCTTCTCGAGATCTCTGCCAACAATGATGTGTGATACCA	962
Dp	945	CAGTGATATGAGACGAGTGCAGCTTCTCGAGATCTCTGCCAACAATGATGTGTGATACCA	1004
Qy	963	GCCGGGCACATACTTCTGCTCCTCGCTCCACAGGCTACATCTCTGTGATGCAACACGAAG	1022
Dp	1005	GCCGGGCACATACTTCTGCTCCTCGCTCCACAGGCTACATCTCTGTGATGCAACACGAAG	1064
Qy	1023	CTGCCAAGACATCAAGAAATGTGAGCAAGGAAACCAACAGGTGCAACCTCGACAGCAAGCTG	1082
Dp	1065	CTGCCAAGACATCAAGAAATGTGAGCAAGGAAACCAACAGGTGCAACCTCGACAGCAAGCTG	1124
Qy	1083	CTACAAATTTACAAGGGGGCTTCAAAATGCATGACCCCAATCCGCTGTGAGAGCCCTTAAT	1142
Dp	1125	CTACAAATTTACAAGGGGGCTTCAAAATGCATGACCCCAATCCGCTGTGAGAGCCCTTAAT	1184
Qy	1143	GAGATCGATATAACCGCTGATGTGCTCGTGCAGAAACCTCGTGCTCAAGACACAGCC	1202
Dp	1185	GAGATCGATATAACCGCTGATGTGCTCGTGCAGAAACCTCGTGCTCAAGACACAGCC	1244
Qy	1203	CTTTACCATCTTGTACCAGGACATGAGAGTGTGTCAGAGAGGCTCCGTTCCGGCTGACAT	1262
Dp	1245	CTTTACCATCTTGTACCAGGACATGAGAGTGTGTCAGAGAGGCTCCGTTCCGGCTGACAT	1304
Qy	1263	CTTCCAAATGCAACCCACACAGCCCGCTACCTGTGGGGCTATTACATTTTCCAGATCAATC	1322
Dp	1305	CTTCCAAATGCAACCCACACAGCCCGCTACCTGTGGGGCTATTACATTTTCCAGATCAATC	1364
Qy	1323	TGGCAATGAGGGGACAGAAATTTTACATGCGGCAAAAGGGGCCCAATGATGCCACCTGGT	1382
Dp	1365	TGGCAATGAGGGGACAGAAATTTTACATGCGGCAAAAGGGGCCCAATGATGCCACCTGGT	1424
Qy	1383	GATGACACGCCCATCAAAAGGGGCCCGGAAATCAGCTGGACTTGGAAATGATCACTGT	1442
Dp	1425	GATGACACGCCCATCAAAAGGGGCCCGGAAATCAGCTGGACTTGGAAATGATCACTGT	1484
Qy	1443	CAACACGTGATCAAACTTCAAGAGGACAGCTCCGTATCCGATCGAGTGTGGGATATATGTGTGCA	1502
Dp	1485	CAACACGTGATCAAACTTCAAGAGGACAGCTCCGTATCCGATCGAGTGTGGGATATATGTGTGCA	1544
Qy	1503	GTACCCATTTGAGACCTTCGGGCTGTGAGACCTTCGACGCTGCGCTTCATTTGGCACAAGGA	1562
Dp	1545	GTACCCATTTGAGACCTTCGGGCTGTGAGACCTTCGACGCTGCGCTTCATTTGGCACAAGGA	1604
Qy	1563	CAGGAGAAAGAGGAAATTAACAGAGAAATGAGAGCACAACAGCTTTAGGCATTTTCTG	1622

Db	1605	CAGGAGAAGAGAGCAAAATTAACAGAGAGAMWTGAGACCAGMACAAGACTTATGGCATTTTCCTG	1604
QY	1623	CTGACGCTTTCCCGAAGAGTAGCAGCCCCGACTTCTGACTGTACCTGTACTATTG	1678
Db	1665	CTGAACGTTTTCCCAGAGAGATCAAGNCCGACTTCTGTACTCTCACCTGTACTATTG	1720
RESULT 10			
LOCUS	BC006636		
DEFINITION	mus musculus, fibulin 5, clone MGC:5656 IMAGE:3482574, mRNA,	2230 bp	mRNA linear ROD 07-AUG-2002
ACCESSION	BC006636		
VERSION	BC006636.1	GI:13879321	
KEYWORDS	MGC.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.		
REFERENCE	1 (bases 1 to 2230)		
AUTHORS	Straussberg, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-APR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk Email: cgabs@email.nih.gov Tissue Procurement: Lotmar Hennighausen Ph.D., Robin Humphreys CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contract: amg@bcm.tmc.edu Gunnarath, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.		
FEATURES	Source		
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://Image.lnl.gov Series: IRMA Plate: 6 Row: f Column: 8 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6753823. Location/Qualifiers 1..2230 /organism="Mus musculus" /db_xref="taxon:10090" /map="C57BL/6J" /clone="MGC:5656 IMAGE:3482574" /tissue_type="Mammary tumor. WAP-TGF alpha model. 7 months old, gross tissue." /clone_lib="NCI CGAP_Mams" /lab_host="DH10B" /note="Vector: pCMV-SPORT6" 60..1406 /codon_start=1 /product="fibulin 5" /protein_id="AAH06636.1" /db_xref="GI:13879322" /db_xref="locusID:23876" /translation="MPGLKRLITVTTLALMLPHPMNAOQCCTNGFDLRSGQCCLDIDECRTLEPARGDMMCVNONGYLICPTNPYRGPSNPYSTSYSGPPAAPAPVPASNPPTSPLVCRRGYOMDENOGVDCVDVECATDSHCNFTQLCINTEGGYCTGDWYLLLEGOLCIDDECERYGCOOLCANVPGSYCTGNCFETLNDGRSCODVNECEENCVOTCWMTVGSEFCRDGPEGTLEEEDGHCSMDXESFSESLCOHECNMGNSYFGCPGAYVLDDNRSCODINECBHRHTTSTSLDTCYRNLOGSKFKIDISCEPFILITENRCMCFAEHSRCADPFTILLTRMDVVSGRSPADIFOCATTRYGAYTIFIKSSNEGEPYMRQGEPSATLVMTPIKGPRIOIDLLEMITVNTVINFRSSVTRLRIYVSQYPE"		
CDS			
BASE COUNT	562 a	578 c	505 g 585 t

Db	963	GCCTGGCAGATACCTTCTGCTCTCTGCTCCCTCCAGGCTACATCTCTGCTGATGCAACCGAAG	1022
OY	1023	CTGCCAAGACATCAACGAATGTGAGCACAGGACACACGTCGAACCTGCAGCAGACGTG	1082
Db	1023	CTGCCAAGACATCAACGAATGTGAGCACAGGACACACGTCGAACCTGCAGCAGACGTG	1082
OY	1083	CTACAATTTCAAGGGGGCTTCAATGCATGACGCCCATCCGCTGTGAGAGCCTTATATCT	1142
Db	1083	CTACAATTTCAAGGGGGCTTCAATGCATGACGCCCATCCGCTGTGAGAGCCTTATATCT	1142
OY	1143	GAGGATACGATATACCGCTGTATGTGCTCTGCTGAGAACCTTGCTGCAAGACCAAGCC	1202
Db	1143	GAGGATACGATATACCGCTGTATGTGCTCTGCTGAGAACCTTGCTGCAAGACCAAGCC	1202
OY	1203	CTTTACCATCTTGTACCGGGACATGAGAGTGTGTCAAGAGCTTCCTCCGCTGACAT	1262
Db	1203	CTTTACCATCTTGTACCGGGACATGAGAGTGTGTCAAGAGCTTCCTCCGCTGACAT	1262
OY	1263	CTTCCAAATGCACACCCACAGACCCGCTACCCCTGAGGGGCTATATCAATTTTCCAGATCAATC	1322
Db	1263	CTTCCAAATGCACACCCACAGACCCGCTACCCCTGAGGGGCTATATCAATTTTCCAGATCAATC	1322
OY	1323	TGGCAATGAGGGGACAGAAATTTTATCATGTGCGCAAAAGGGGCCCATTCAGTGCACCTGTGT	1382
Db	1323	TGGCAATGAGGGGACAGAAATTTTATCATGTGCGCAAAAGGGGCCCATTCAGTGCACCTGTGT	1382
OY	1383	GATGACAGCGCCCATCAACAAAGGGGCCCGGGAATCAGCTGGACATTTGGAAATGATCACTGT	1442
Db	1383	GATGACAGCGCCCATCAACAAAGGGGCCCGGGAATCAGCTGGACATTTGGAAATGATCACTGT	1442
OY	1443	CAACACTGTCTATCACTTCAGAGGACAGCTCCGTGATCCGACTGCGGATATATGTGTGCGCA	1502
Db	1443	CAACACTGTCTATCACTTCAGAGGACAGCTCCGTGATCCGACTGCGGATATATGTGTGCGCA	1502
OY	1503	GTACCCATTTGAGACCTTCGGGCTGGAGCTTCGAGCGCTCCGCTCATTTGGACCAAGGA	1562
Db	1503	GTACCCATTTGAGACCTTCGGGCTGGAGCTTCGAGCGCTCCGCTCATTTGGACCAAGGA	1562
OY	1563	CAGAGAAGAGAGAAATTAACAGAGAGATGAGAGCGACACAGACGCTTAGGCATTTCTGT	1622
Db	1563	CAGAGAAGAGAGAAATTAACAGAGAGATGAGAGCGACACAGACGCTTAGGCATTTCTGT	1622
OY	1623	CTGAACGTTTCCCGGAAGAGTCAAGCCCGACTTCTGTACTCTCACTCTGATATTGACAGA	1682
Db	1623	CTGAACGTTTCCCGGAAGAGTCAAGCCCGACTTCTGTACTCTCACTCTGATATTGACAGA	1682
OY	1683	CCTGTCAACCTGCAGAGACTTCCACCCCGCACTTCTGTATGATACAGTTATCAAAAAGTAT	1742
Db	1683	CCTGTCAACCTGCAGAGACTTCCACCCCGCACTTCTGTATGATACAGTTATCAAAAAGTAT	1742
OY	1743	ATCATTTGCTCCCTGATAGAAAGATTGTGTGAATTTTCAAGGGCTTCAGTTTATTTCCA	1802
Db	1743	ATCATTTGCTCCCTGATAGAAAGATTGTGTGAATTTTCAAGGGCTTCAGTTTATTTCCA	1802
OY	1803	CTATTTTCAAGAAATATAGATTAGTTTGGGGGGGTGAGCTATATGTCAAAACAGTGTG	1862
Db	1803	CTATTTTCAAGAAATATAGATTAGTTTGGGGGGGTGAGCTATATGTCAAAACAGTGTG	1862
OY	1863	AACAGCTTGTCTACATCTTTCACCTCTTCCACTCTTCTCTCACTGTTACTGCTTTG	1922
Db	1863	AACAGCTTGTCTACATCTTTCACCTCTTCCACTCTTCTCTCACTGTTACTGCTTTG	1922
OY	1923	CAAGAGCCC -GGAGAGCTGGCGGGGAA - -CCCTGGGAGTAGTATTTGCTTTTGGCTAC	1979
Db	1923	CAAGAGCCC -GGAGAGCTGGCGGGGAA - -CCCTGGGAGTAGTATTTGCTTTTGGCTAC	1982
OY	1980	ACAG -AGAAGGCTATGTAACCAAAACCAAGCAGAGATC	2015
Db	1983	ACAGAGAGGCTATGTAACCAAAACCAAGCAGAGATC	2019

LOCUS	A84086	1720 bp	DNA	linear	PAT 21-JAN-2000
DEFINITION	Sequence 1 from Patent WO9846746.				
ACCESSION	A84086				
VERSION	A84086.1	GI:6733224			
KEYWORDS					
SOURCE	unidentified.				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 1720)				
AUTHORS	Li, H. and Olsen, H.S.				
TITLE	EXTRACELLULAR/EPIDERMAL GROWTH FACTOR LIKE PROTEIN				
JOURNAL	Patent: WO 9846746-A 1 22-OCT-1998;				
	HUMAN GENOME SCIENCES INC (US); LI HAODONG (US)				
FEATURES	Location/Qualifiers				
source	1..1720				
	/organism="unidentified"				
	/db_xref="taxon:32644"				
	211..1557				
CDS	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAB69285.1"				
	/db_xref="GI:6733225"				
	/translation="MPGIRILIVTILALCLPSPGNAOCTNGFPLDROSGLCLDIDIECIRIPEACRGDMCVNONGGYLICIPRNPYRGPYSNPSTPYSGYPAAPLSAPNYPRIISRLICRGFYOMDESNOCYDVCAIDSHQCNPTQICITETGEGSCIDGYWNLBEOCLDIDECRYGCOQLCANYPGSSCNPGFTLINEDGRSCOVNECATENPCVOTVNTYGSFICRCDDPGYELDEDEHGSYDCEGSEFSEFICHECVNPGYFCSGPGYIILDDNRSODDINCEHRNHTCNLOOTCVNLOGQFPCIDIREEPILRISDNRMCKPAENPGCHDOPFTLLYRMDVSVRSRPADIFQMDATTRIRGATYIITQIKSGNREFYMRQTPISATLVMTIRPLKPRFRIQDLEMTVNTVINFSGSVIRLITVISOIP"				
sig.peptide	211..284				
mat.peptide	287..1554				
BASE COUNT	396 a 515 c 434 g 372 t 3 others				
ORIGIN					
Query Match	71.9% Score 1674.2; DB 6; Length 1720;				
Best Local Similarity	99.8%; Pred. No. 0;				
Matches 1673; Conservative	2; Mismatches 1; Indels 0; Gaps 0;				
QY	3	CCCGGCGGTCACCCGCTGCTCTCCACAGATCGTCGGCCCTCTGGATATAACACC	62		
DB	45	CCCGGCGGTCACCCGCTCTCTCCACAGATCGTCGGCCCTCTGGATATAACACC	104		
QY	63	GCGAGCCCGGAGGGCCAGAGAGGGCGAGCTGCCGAGCTCTCGGGGCTCCGCCG	122		
DB	105	GCGAGCCCGGAGGGCCAGAGAGGGCCAGCTGCCGAGCTCTCGGGGCTCCGCCG	164		
QY	123	CGAGCTTCTCTGGCCTTGCACATCTCTCTCGGGGCTTGAGACATGCAGGATATA	182		
DB	165	CGAGCTTCTCTGGCCTTGCACATCTCTCTCGGGGCTTGAGACATGCAGGATATA	224		
QY	183	AAGATATCTACTGTTTACATCTTGGGCTCTGTCTTCCAAAGCCCTGGGAATGACA	242		
DB	225	AAGATATCTACTGTTTACATCTTGGGCTCTGTCTTCCAAAGCCCTGGGAATGACA	284		
QY	243	ACAGTGCACGAATGGCTTTGACATCGTGATCGGCAGTCAAGTAAGTTAATTAAGA	302		
DB	285	ACAGTGCACGAATGGCTTTGACATCGTGATCGGCAGTCAAGTAAGTTAATTAAGA	344		
QY	303	ATGCCGAACATCCCGGAGGCTCGCCGAGAGACATGATGTGTGTTAACCAAAATGCGG	362		
DB	345	ATGCCGAACATCCCGGAGGCTCGCCGAGAGACATGATGTGTGTTAACCAAAATGCGG	404		
QY	363	GTATTTATGATTTCCCGGAGCAAAACCTGTGTATGAGGGCCCTACTGAAACCCCTACTC	422		
DB	405	GTATTTATGATTTCCCGGAGCAAAACCTGTGTATGAGGGCCCTACTGAAACCCCTACTC	464		
QY	423	GACCCCTACACAGGTGCGTACCGACAGAGCGGCCACACTCTCAGTCCAAACTATAC	482		
DB	465	GACCCCTACACAGGTGCGTACCGACAGAGCGGCCACACTCTCAGTCCAAACTATAC	524		
QY	483	CACGATCTCCAGGCTCTTATATGCGCGGTTTGATATACAGATGATGAACCAACATATG	542		

Db	1501	IGTACTATTGCAAGACCTGTCAACCTGCGAGACCTGCCAACCCCAAGTTCCTATGATACAGT	1560
Oy	1729	TATCAAAAAGTATTATCATCTTGTCTCCCTCGATAGAGATTTGTTGGTAATTTTCAAGCCT	1788
Db	1561	TATCAAAAAGTATTATCATCTTGTCTCCCTCGATAGAGATTTGTTGGTAATTTTCAAGCCT	1620
Oy	1789	TCAGTTTATTTCCACTATTTTTCAAAGAAAATAGATTAGTTGGTGGGGGCGTGCAGCTAT	1848
Db	1621	TCAGTTTATTTCCACTATTTTTCAAAGAAAATAGATTAGTTGGTGGGGGCGTGCAGCTAT	1680
Oy	1849	GTTCAAAAGACTGTGAAACAGCTTGTCTGTCACTTCTTCACCTCTTCACCTCTTCTCACT	1908
Db	1681	GTTCAAAAGACTGTGAAACAGCTTGTCTGTCACTTCTTCACCTCTTCACCTCTTCTCACT	1740
Oy	1909	GTTGTATACGCTTTCCAAAGAACCCGGGAGCTGGGGGGAACCCCTGGAGATGACTAGTTGC	1968
Db	1741	GNGTACTGCTTTCCAAAGAACCCGGGAGCTGGGGGGAACCCCTGGAGATGACTAGTTAGC	1800
Oy	1969	TTTTTGGCTACACAGAGAGAGCTATGTAAACAAACACACAGCAGATCGAAGGGTTTTTAC	2028
Db	1801	TTTTTGGCTACACAGAGAGAGCTATGTAAACAAACACACAGCAGATCGAAGGGTTTTTAC	1860
Oy	2029	AGATGAGTTTCCAAACCAACATGCGGATATTTTCAACCATAAAGATTCAGTTGTCCT	2088
Db	1861	AGATGAGTTTCCAAACCAACATGCGGATATTTTCAACCATAAAGATTCAGTTGTCCT	1920
Oy	2089	TAAATTTGTATTAACGGTTTAAATTCGTGCTGTTGATTTTGAATTTTAAAAAATATGT	2148
Db	1921	TAAATTTGTATTAACGGTTTAAATTCGTGCTGTTGATTTTGAATTTTAAAAAATATGT	1980
Oy	2149	CGTAGAATTCCTTGAAAAGGCTTCACACACATCTCTATG	2187
Db	1981	CGTAGAATTCCTTGAAAAGGCTTCACACACATCTCTATG	2019

FEATURES	Source
LOCUS	AF093118
DEFINITION	Homo sapiens UP50 mRNA, complete cds.
ACCESSION	AF093118
VERSION	AF093118.1
KEYWORDS	GI:3676821
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 2019)
TITLE	Zemel, R., Sholto, O. and Shaul, Y.
JOURNAL	Direct Submission
	Submitted (15-SEP-1998) Dept. of Molecular Genetics, Weizmann Institute of Science, Rehovot 76100, Israel
	Location/Qualifiers
	1..2019
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/tissue_type="urine"
	1..168
	169..1515
	/note="50KD; similar to human SI-5 encoded by Genbank accession Number U03877; possible membrane and extracellular matrix protein; urine p50 protein"
	/codon_start=1
	/product="up50"
	/protein_id="AAC62107.1"
	/db_xref="GI:3676822"
	/translation="MPGIRKILVYTLIALCLPSPGNAOACITNGFDLDRSGOCLIDECRTIPACRGDMCMVCONGGLYCHSRSTNIVYRGPNSTPSPGYPAAPPLSAPNYPISRLICRFEGYOMDESNOCVDECATDSHOCNPVICINMGKGYCTSDGYWLTGGQCLIDECRGYCOOLCANVPGSISCTNGPFLINDGSCODVNCATENPCVOTCVNTYGSFTICRCRDPYELKEDGVHSHDDECSFSEFLQHECVNPDTYFCSCPGTYLIDDDNSCODINECHRNHTCLQOTCYNLIDGFCIDPIKEEPEYLAISDRCKCAPENPGSRDQPTFLVMDMVSGRSVADIFQWATTRVGAAYIQLQISGNGREFYMRGTGISLTWLRPIKSPREIOLDLEITVNTVTVNFRGSSVIRLRIVSYPP"

[illegible]


```
QY 303 ATGCCAGCAATCCCGAGGCGTCGAGAGACATGATGTGTAAACCAAAATGGCGG 362
DB |||||||
DB 318 ATGCCAGCAATCCCGAGGCGTCGAGAGACATGATGTGTAAACCAAAATGGCGG 377
QY 363 GATATTATGCAATTCCTCCGAGCAAAACCTGTGTATCGAGGGCCCTACTCGAACCCCTACTC 422
DB |||||||
DB 378 GATATTATGCAATTCCTCCGAGCAAAACCTGTGTATCGAGGGCCCTACTCGAACCCCTACTC 437
QY 423 GACCCCTACTAGGTCCTGACCAAGACAGTGTGCCCAACCACTCTACCTCCAACTATCC 482
DB |||||||
DB 438 GACCCCTACTAGGTCCTGACCAAGACAGTGTGCCCAACCACTCTACCTCCAACTATCC 497
QY 483 CACGATCTCAGGCGCTTATATGCGCGCTTGGATACAGATGGATGAAGCAACGATG 542
DB |||||||
DB 498 CACGATCTCAGGCGCTTATATGCGCGCTTGGATACAGATGGATGAAGCAACGATG 557
QY 543 TGTGATGTGACAGAGTGTGCAACAGATTCCACAGATGCAACCCCAACCCAGATCTGCAT 602
DB |||||||
DB 558 TGTGATGTGACAGAGTGTGCAACAGATTCCACAGATGCAACCCCAACCCAGATCTGCAT 617
QY 603 CAATACTGAAGGCGGCTACACCGTCCTGACAGAGATATTGGCTTCTGGAAAGCCA 662
DB |||||||
DB 618 CAATACTGAAGGCGGCTACACCGTCCTGACAGAGATATTGGCTTCTGGAAAGCCA 677
QY 663 GTGCTTAGACATTTGATGAAATGTGCTATGTGTTACTGACAGAGTCTGTGGAATGTTC 722
DB |||||||
DB 678 GTGCTTAGACATTTGATGAAATGTGCTATGTGTTACTGACAGAGTCTGTGGAATGTTC 737
QY 723 TGGATCTTATTTCTTGTACATGCAACCCCTGTGTTTACCTCAATGAGAGTGAAGTCTTG 782
DB |||||||
DB 738 TGGATCTTATTTCTTGTACATGCAACCCCTGTGTTTACCTCAATGAGAGTGAAGTCTTG 797
QY 783 CCAAGATGTGAAGAGTGTGACACGAGAAACCCCTGTCGCAAAACCTGCTCAACACTA 842
DB |||||||
DB 798 CCAAGATGTGAAGAGTGTGACACGAGAAACCCCTGTCGCAAAACCTGCTCAACACTA 857
QY 843 CGGCTCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAAGAGTGGCTTCATTG 902
DB |||||||
DB 858 CGGCTCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAAGAGTGGCTTCATTG 917
QY 903 CAGGATATGAGAGAGTGTGACAGTCTGTGAGTTCCTGTGCAACATGAGTGTGAACCA 962
DB |||||||
DB 918 CAGGATATGAGAGAGTGTGACAGTCTGTGAGTTCCTGTGCAACATGAGTGTGAACCA 977
QY 963 GCCCGGCAATCTCTGCTCTGCGCTCAGAGCTCATCTGCTGGATGACAAACGAG 1022
DB |||||||
DB 978 GCCCGGCAATCTCTGCTCTGCGCTCAGAGCTCATCTGCTGGATGACAAACGAG 1037
QY 1023 CTGCGAAGACATCAACGAATGTGACACAGAACCAACACGTGCACAGCAGAGCTG 1082
DB |||||||
DB 1038 CTGCGAAGACATCAACGAATGTGACACAGAACCAACACGTGCACAGCAGAGCTG 1097
QY 1083 CTACAAATTTACAAAGGGGCTTCAAAATGATGATGACCCCAACCCCTGTGAGAGAGCTTATCT 1142
DB |||||||
DB 1098 CTACAAATTTACAAAGGGGCTTCAAAATGATGATGACCCCAACCCCTGTGAGAGAGCTTATCT 1157
QY 1143 GAGGATCAGTATTAACCGCTGTATGTCTGCTGAGAACCCCTGTGAGAGAGCAGCC 1202
DB |||||||
DB 1158 GAGGATCAGTATTAACCGCTGTATGTCTGCTGAGAACCCCTGTGAGAGAGCAGCC 1217
QY 1203 CTTTACCATCTTTGTACCGGGAGACAGTGTGTGAGAGAGCTCCGTTCCGCTGACAT 1262
DB |||||||
DB 1218 CTTTACCATCTTTGTACCGGGAGACAGTGTGTGAGAGAGCTCCGTTCCGCTGACAT 1277
QY 1263 CTTTCAATGTGACGACGACCGCTACCCGCGGCTTATTCATTTTCCATGTCAAATC 1322
DB |||||||
DB 1278 CTTTCAATGTGACGACGACCGCTACCCGCGGCTTATTCATTTTCCATGTCAAATC 1337
QY 1323 TGGGAATGAGGGGAGAGAAATTTTACATGCGGCAAAAGGGCCCATCAGTGGCCACCTGGT 1382
DB |||||||
DB 1338 TGGGAATGAGGGGAGAGAAATTTTACATGCGGCAAAAGGGCCCATCAGTGGCCACCTGGT 1397
QY 1383 GATGACACGCGCCCATCAAAAGGGCCCGGGAATTCAGCTGTGAAATGATCACTGT 1442

|||||
DB 1398 GATGACACGCGCCCATCAAAAGGGCCCGGGAATTCAGCTGTGAAATGATCACTGT 1457
QY 1443 CAACACTGTATCAACTTCAGAGAGAGCTCCGCTGATCCGACTGGGATATATGTGCGCA 1502
DB |||||||
DB 1458 CAACACTGTATCAACTTCAGAGAGAGCTCCGCTGATCCGACTGGGATATATGTGCGCA 1517
QY 1503 GTACCCATCTGAGCTCGGGCTGAGAGCTCCGAGAGCTGCTCATTTGGCAGCAAGGGA 1562
DB |||||||
DB 1518 GTACCCATCTGAGCTCGGGCTGAGAGCTCCGAGAGCTGCTCATTTGGCAGCAAGGGA 1577
QY 1563 CAGGAGAAAGAGGAAATTAACAGAGATGAGAGGACACAGAGCTTAAAGCATTTCTG 1622
DB |||||||
DB 1578 CAGGAGAAAGAGGAAATTAACAGAGATGAGAGGACACAGAGCTTAAAGCATTTCTG 1637
QY 1623 CTGAAGCTTCCCGGAAGAGTGTGAGCCCGCTTCTGACCTCAACCTGATCTATTTGAGA 1662
DB |||||||
DB 1638 CTGAAGCTTCCCGGAAGAGTGTGAGCCCGCTTCTGACCTCAACCTGATCTATTTGAGA 1697
QY 1683 CCTGTACCCCTGAGAGCTTGCACCCCAAGTTCTATGATACAGTATCAAAAAGTAT 1742
DB |||||||
DB 1698 CCTGTACCCCTGAGAGCTTGCACCCCAAGTTCTATGATACAGTATCAAAAAGTAT 1757
QY 1743 ATCATTTGCTCCCTGATGAGAAATGTGTGTAATTTCAAGGCTTCAAGTATTTTCCA 1802
DB |||||||
DB 1758 ATCATTTGCTCCCTGATGAGAAATGTGTGTAATTTCAAGGCTTCAAGTATTTTCCA 1817
QY 1803 CTATTTTCAAAAGAAATAGATTAGTTTCCGAGGCTGTGAGTATGATTAAGAGACGTG 1862
DB |||||||
DB 1818 CTATTTTCAAAAGAAATAGATTAGTTTCCGAGGCTGTGAGTATGATTAAGAGACGTG 1877
QY 1863 AACAGCTGTCTACCTTCTTACCTCTTCCACTCTCTCTGACAGTGTGTTACTGCTTG 1922
DB |||||||
DB 1878 AACAGCTGTCTACCTTCTTACCTCTTCCACTCTCTCTGACAGTGTGTTACTGCTTG 1937
QY 1923 CAAAGACCCGGGAGGCTGGGGGAAACCCGAGGAGTGTGATGTTTGGCTTACACA 1982
DB |||||||
DB 1938 CAAAGACCCGGGAGGCTGGGGGAAACCCGAGGAGTGTGATGTTTGGCTTACACA 1997
QY 1983 GAGAAGGCTATGTAACAAACCAACACAGACAGATCGAAGGTTTGTAGAGATGTGTTCAA 2042
DB |||||||
DB 1998 GAGAAGGCTATGTAACAAACCAACACAGACAGATCGAAGGTTTGTAGAGATGTGTTCAA 2057
QY 2043 AACATGCTGTGATTTTCAACATTAAGAAAGTTCAAGTGTCTTAAATTTGTATPAC 2102
DB |||||||
DB 2058 AACATGCTGTGATTTTCAACATTAAGAAAGTTCAAGTGTCTTAAATTTGTATPAC 2117
QY 2103 GGTTAATTTCTGTGTTTCAATTTTGTAGATTTTAAAAAATGTGTAAGATTTGCTTC 2162
DB |||||||
DB 2118 GGTTAATTTCTGTGTTTCAATTTTGTAGATTTTAAAAAATGTGTAAGATTTGCTTC 2177
QY 2163 GAAAGGCTTACAGACATGCTATGTCTGTCTCCAAACCAACGAGTCTCTCATATTT 2222
DB |||||||
DB 2178 GAAAGGCTTACAGACATGCTATGTCTGTCTCCAAACCAACGAGTCTCTCATATTT 2237
QY 2223 AGCCAGAGTCTTCTTGTAGAGAGCCCTTAATCTGCTTCTTAAATTTTAAATTTTACC 2282
DB |||||||
DB 2238 AGCCAGAGTCTTCTTGTAGAGAGCCCTTAATCTGCTTCTTAAATTTTAAATTTTACC 2297
QY 2283 GGATTTGAGATCAGAGGTCTCCAAACGATTAATTAATTTGAAGAGA 2328
DB |||||||
DB 2298 GGATTTGAGATCAGAGGTCTCCAAACGATTAATTAATTTGAAGAAA 2343

RESULT 7
HSA133490 HSA133490 2019 bp mRNA linear PRI 11-MAR-1999
LOCUS Homo sapiens fibulin-5.
DEFINITION A133490
ACCESSION A133490.1 GI:4490529
VERSION A133490.1
KEYWORDS FTBL-5 gene; fibulin-5.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
```

```

OY 1443 CAACATGTCATCAATTCAGAGCAGCTCCGATCCGATCGCATATATGTGTCCA 1502
DB 1724 CAACATGTCATCAATTCAGAGCAGCTCCGATCCGATCGCATATATGTGTCCA 1783
OY 1503 GTACCATTTCTAGAGCTGGGCTGGAGCCTCCGACGCTGCTCTCATTTGGCACCAGGGA 1562
DB 1784 GTACCATTTCTAGAGCTGGGCTGGAGCCTCCGACGCTGCTCTCATTTGGCACCAGGGA 1843
OY 1563 CAGAGAGAGAGAGAAATTAACAGAGAAATGAGAGGACACAGACGTTAGCATTTCTCTG 1622
DB 1844 CAGAGAGAGAGAGAAATTAACAGAGAAATGAGAGGACACAGACGTTAGCATTTCTCTG 1903
OY 1623 CTGAACGTTTCCCGAGAGAGTCCGACCTCTGACTCTCAGCTGACTATTTGACGA 1682
DB 1904 CTGAACGTTTCCCGAGAGAGTCCGACCTCTGACTCTCAGCTGACTATTTGACGA 1963
OY 1683 CCTGTACCCCTCAGAGACTTGGCCACCCCAAGTTCTATGATACAGTTATCAAAAAGTAT 1742
DB 1964 CCTGTACCCCTCAGAGACTTGGCCACCCCAAGTTCTATGATACAGTTATCAAAAAGTAT 2023
OY 1743 ATCATTTGCTCCCTGATAGAGATTTGTTGTAATTTTCAAGGCTTCAGTTATTTTCCA 1802
DB 2024 ATCATTTGCTCCCTGATAGAGATTTGTTGTAATTTTCAAGGCTTCAGTTATTTTCCA 2083
OY 1803 CTATTTTCAAGAGAAATAGATTAGTTTGGGGGGGCTGAGTCTATGTTTCAAGAGACTGTG 1862
DB 2084 CTATTTTCAAGAGAAATAGATTAGTTTGGGGGGGCTGAGTCTATGTTTCAAGAGACTGTG 2143
OY 1863 AACAGCTTGTCTGCTACCTTTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 1922
DB 2144 AACAGCTTGTCTGCTACCTTTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 2203
OY 1923 CAAGAGCCGGAGAGCTGGGGGGGACCTGGGAGTACTGTTGTTTGGGTTTGGGTTTGGGTTT 1982
DB 2204 CAAGAGCCGGAGAGCTGGGGGGGACCTGGGAGTACTGTTGTTTGGGTTTGGGTTTGGGTTT 2263
OY 1983 GAGAGAGCTATGTAAACAAACACACAGAGATGAGAGGTTTGTAGAGATGTGTTCAA 2042
DB 2264 GAGAGAGCTATGTAAACAAACACACAGAGATGAGAGGTTTGTAGAGATGTGTTCAA 2323
OY 2043 AACCATGCTGCTATTTTCAACCATTAAGAAAGTTTGTAGTTTGTCTTAAATTTGTATAC 2102
DB 2324 AACCATGCTGCTATTTTCAACCATTAAGAAAGTTTGTAGTTTGTCTTAAATTTGTATAC 2383
OY 2103 GATTAAATCTGCTCTGCTATTTTGAATTTTAAATAATATGCTAGTAATTCCTTC 2162
DB 2384 GATTAAATCTGCTCTGCTATTTTGAATTTTAAATAATATGCTAGTAATTCCTTC 2443
OY 2163 GAAAGGCTTTCAGACATGCTATGTGTCTCTCCAAACCCAGTCTCTCTCATTTT 2222
DB 2444 GAAAGGCTTTCAGACATGCTATGTGTCTCTCCAAACCCAGTCTCTCTCATTTT 2503
OY 2223 AGCCAGTGTCTTCTTGAAGACCCCTTAATCTTCTCTTCTTGAATTTTACCCTAAT 2282
DB 2504 AGCCAGTGTCTTCTTGAAGACCCCTTAATCTTCTCTTCTTGAATTTTACCCTAAT 2563
OY 2283 GGATTTGGAATGAGAGGCTCCCAACGATTAATATTTTGAAGAGA 2328
DB 2564 GGATTTGGAATGAGAGGCTCCCAACGATTAATATTTTGAAGAGA 2609

```

```

RESULT 6
BC022280 2367 bp mRNA linear PRI 04-FEB-2002
LOCUS Homo sapiens, clone MGC:22412 IMAGE:4693953, mRNA, complete cds.
ACCESSION BC022280.1 GI:18490144
VERSION MGC.
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 2367)

```

AUTHORS

Strausberg, R.
Direct Submission
Submitted (01-FEB-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaps-remail.nih.gov

COMMENT

Tissue Procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mdickson@stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAL Plate: 36 Row: n Column: 23
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 5453649.

FEATURES

source

1. 2367
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:22412 IMAGE:4693953"
/tissue_type="Lung"
/clone_id="NIH_MGC_77"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
184. 1530
/codon_start=1
/product="Unknown (protein for MGC:22412)"
/protein_id="AAH22280.1"
/db_xref="gi:18490145"
/translation="MGIKRLITVTLALCLSPGNAOAGCTNGFDLRSGGCLDID
ECRTPEACGDMCVNNGCYLCTPRNPYRPPXSPYSPYGPAPAPASAP
NPTFISPLICRFGRYQDNESNOCVADVEACDTHOACNFOICINTEGGTCTGTYW
LIEGQCLIDIECRGYCOQLCANVPSYSTCYNEDGRSCODVNECACTENPCV
OTCVNTRYSFICRCDPELEEDVHSDMECSFELCOHECVNOGTYVSCDPG
YILDDNRSCODINECHRNHTCMLOOTCYMLQGFKIDIRCEPEYLRISDNCMC
PANPGCRDPTILYRDMDVVSGRSPADIFQOATRYGAYYIFQIKSGNEERF
YMOGTGSLITVMTIRPKIPREIOLDLEMTIVTVINFRSSVIRLRIYVQYF"

CDS

BASE COUNT

588 a 650 c 547 g 582 t

ORIGIN

Query Match 99.7% Score 2321.2; DB 9; Length 2367;
Best Local Similarity 99.9% Pred. No. 0;
Matches 2323; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

OY 3 CCGGGGCTCTCCCGGTGCTCTCAGACAGCTGCGGCCCTCTGGAATAAACAACC 62
DB 18 CCCGGGCTCTCCCGGTGCTCTCAGACAGCTGCGGCCCTCTGGAATAAACAACC 77
OY 63 GCGAGCCCGGAGGGCCAGAGAGAGCGAGCTGCCGAGCTCTCCGGGGGTCGCCGCC 122
DB 78 GCGAGCCCGGAGGGCCAGAGAGAGCGAGCTGCCGAGCTCTCCGGGGGTCGCCGCC 137
OY 123 CGAGCTTCTTCGCGCTTCGATCTCTCTCGCGGCTTGTGACATGCCAGATAAA 182
DB 138 CGAGCTTCTTCGCGCTTCGATCTCTCTCGCGGCTTGTGACATGCCAGATAAA 197
OY 183 AAGGATACACAGTTACATTCGCTCTCTCTCTCAAGCCCTGGGAATGACAGGC 242
DB 198 AAGGATACACAGTTACATTCGCTCTCTCTCTCAAGCCCTGGGAATGACAGGC 257
OY 243 ACAGTCACAGAAAGGCTTACCTGATTCGCAATGCGACATGAGAGAGTATATGATGA 302
DB 258 ACAGTCACAGAAAGGCTTACCTGATTCGCAATGCGACATGAGAGAGTATATGATGA 317

```

QY 2103 GCTTAATCTGCTTGTGATTTGATTTTAAATAATATGCTAGAAATCTTC 2162
| | | | |
Db 2384 GGTTAATCTGCTTGTGATTTGATTTTAAATAATATGCTAGAAATCTTC 2443
| | | | |
QY 2163 GAAAGCCCTTCAGACATGCTATGTTCTGCTCCAAACAGTCTCTCCATTT 2222
| | | | |
Db 2444 GAAAGCCCTTCAGACATGCTATGTTCTGCTCCAAACAGTCTCTCCATTT 2503
| | | | |
QY 2223 AGCCAGTCTTCTTGTGAGAGCCCTTAATCTGCTTCTTGTAGATTTTACCATT 2282
| | | | |
Db 2504 AGCCAGTCTTCTTGTGAGAGCCCTTAATCTGCTTCTTGTAGATTTTACCATT 2563
| | | | |
QY 2283 GGATTGGAATGCAGAGCTTCGCAATGATTAATATTGAAGA 2328
| | | | |
Db 2564 GGATTGGAATGCAGAGCTTCGCAATGATTAATATTGAAGA 2609
| | | | |
RESULT 5
LOCUS AX403659 2609 bp DNA Linear PAT 14-JUN-2002
DEFINITION Sequence 14 from Patent WO0077037.
ACCESSION AX403659
VERSION AX403659.1 GI:21437116
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1
Ashkenazi, A., Baker, K., Botstein, D., Desnovers, L., Eaton, D.L.,
Ferreira, N., Fong, W.Q., Gao, W.Q., Gerber, H., Gerlitz, M.E.,
Goddard, A., Godowski, P., Gurney, A., Kijavlin, I.J., Mather, J.,
Napier, M., Pan, J., Paoni, N., Roy, M., Tamas, D., Watanabe, C.,
Williams, P.M., Wood, W.I. and Zhang, Z.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
the same
JOURNAL Patent: WO 0077037-A 14 21-DEC-2000;
Genentech Inc. (US)
FEATURES
source 1.2609
location/Qualifiers
BASE COUNT 622 a 718 c 622 g 647 t
ORIGIN
Query Match 99.88: Score 2322.8; DB 6; Length 2609;
Best Local Similarity 99.98; Pred. No. 0;
Matches 2324; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 363 GATTATGCAATCCCGGACAAACCTGTGTATGAGAGGCGCTACTGAAACCCCTACTC 422
| | | | |
Db 644 GTATTATGCAATCCCGGACAAACCTGTGTATGAGAGGCGCTACTGAAACCCCTACTC 703
| | | | |
QY 423 GACCCCTACTCAGTCCGATACCGAGAGCTGCCCAACACTCTCAGCTCCAAATATCC 482
| | | | |
Db 704 GACCCCTACTCAGTCCGATACCGAGAGCTGCCCAACACTCTCAGCTCCAAATATCC 763
| | | | |
QY 483 CAGATCTCAGGCTCTTATATGCGCTTGTGATACCAATGATGAAACCAATATG 542
| | | | |
Db 764 CAGATCTCAGGCTCTTATATGCGCTTGTGATACCAATGATGAAACCAATATG 823
| | | | |
QY 543 TGTGATGTTGAGAGGTTGCAACAGATTCACACAGTGCAGCCCAACCCAGATCTGCAT 602
| | | | |
Db 824 TGTGATGTTGAGAGGTTGCAACAGATTCACACAGTGCAGCCCAACCCAGATCTGCAT 883
| | | | |
QY 603 CAATCTGAAGGGGTTACACTGCTCTGCACCGAGAGATTTGGCTTCTGAAGGCCA 662
| | | | |
Db 884 CAATCTGAAGGGGTTACACTGCTCTGCACCGAGAGATTTGGCTTCTGAAGGCCA 943
| | | | |
QY 663 GTGCTTAGACATGATGATGATGCTATGTTGTTAGTGCAGAGCTCTGTGGAAATGTTCC 722
| | | | |
Db 944 GTGCTTAGACATGATGATGATGCTATGTTGTTAGTGCAGAGCTCTGTGGAAATGTTCC 1003
| | | | |
QY 723 TGGATCTATTTCTTGTACATGCAACCTGTTTACCTCAATGAGATGGAAGGCTTG 782
| | | | |
Db 1004 TGGATCTATTTCTTGTACATGCAACCTGTTTACCTCAATGAGATGGAAGGCTTG 1063
| | | | |
QY 783 CCAAGATGTGAACAGAGTGTCCACCGAACCCTGCTGCAACCTGCTCAACACCTA 842
| | | | |
Db 1064 CCAAGATGTGAACAGAGTGTCCACCGAACCCTGCTGCAACCTGCTCAACACCTA 1123
| | | | |
QY 843 CGGCTCTTCACTGCGCGGTGACCCAGATATGAACTGAGAAAGTGGCGTTCAATG 902
| | | | |
Db 1124 CGGCTCTTCACTGCGCGGTGACCCAGATATGAACTGAGAAAGTGGCGTTCAATG 1183
| | | | |
QY 903 CAGTGTATGAGAGAGTGCAGCTTCTGATGTTCTGTCGCAACATGATGTTGAACCA 962
| | | | |
Db 1184 CAGTGTATGAGAGAGTGCAGCTTCTGATGTTCTGTCGCAACATGATGTTGAACCA 1243
| | | | |
QY 963 GCCCGCACACTTCTGCTCTGCGCTCCAGAGCTACATCTGCTGATGCAACCGAG 1022
| | | | |
Db 1244 GCCCGCACACTTCTGCTCTGCGCTCCAGAGCTACATCTGCTGATGCAACCGAG 1303
| | | | |
QY 1023 CTGCGCAAGACATCAAGAAATGTAGACAGGAACACACGTCGACCGAGAGCTG 1082
| | | | |
Db 1304 CTGCGCAAGACATCAAGAAATGTAGACAGGAACACACGTCGACCGAGAGCTG 1363
| | | | |
QY 1083 CTACAAATTTCAAGAGGGGCTTCAAAATGCATGACCCCATCGCTGTGAGAGCTTATCT 1142
| | | | |
Db 1364 CTACAAATTTCAAGAGGGGCTTCAAAATGCATGACCCCATCGCTGTGAGAGCTTATCT 1423
| | | | |
QY 1143 GAGGATCAGTGAATCCGCTGTATGTCTCTGCTGAGAACCTGTGCTGACAGACAGCC 1202
| | | | |
Db 1424 GAGGATCAGTGAATCCGCTGTATGTCTCTGCTGAGAACCTGTGCTGACAGACAGCC 1483
| | | | |
QY 1203 CTTTACCATCTTACCGGGAGCATGAGAGTGTGACAGAGCTCCGTTCCGCTGACAT 1262
| | | | |
Db 1484 CTTTACCATCTTACCGGGAGCATGAGAGTGTGACAGAGCTCCGTTCCGCTGACAT 1543
| | | | |
QY 1263 CTTTCAAAATGCAAGCAGACCGCTACCGTGGGGCTTATTAATTTTCCAGATCAATTC 1322
| | | | |
Db 1544 CTTTCAAAATGCAAGCAGACCGCTACCGTGGGGCTTATTAATTTTCCAGATCAATTC 1603
| | | | |
QY 1323 TGGGAATGAGGGCAGAGAAATTTTACATGCGGCAAGGGGCCCATCACTGACCTCTGCT 1382
| | | | |
Db 1604 TGGGAATGAGGGCAGAGAAATTTTACATGCGGCAAGGGGCCCATCACTGACCTCTGCT 1663
| | | | |
QY 1383 GATGACAGCGCCCATCAAGAGGGGCCGGAATTCACCTGACCTTGAATGATCACCTGT 1442
| | | | |
Db 1664 GATGACAGCGCCCATCAAGAGGGGCCGGAATTCACCTGACCTTGAATGATCACCTGT 1723
| | | | |

BASE COUNT	622 a	718 c	622 g	647 t
ORIGIN				

Query Match	99.88;	Score 23322.8;	DB 6;	Length 2609;
Best Local Similarity	99.98;	Pred. NO. 0;		
Matches 2324; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0

QY	3	CCGGGCGCTCCCCGGTGCCTTCACAGATCGGTGGGCCCTTGGAAATTAACACC	62
Db	284	CCGGGCGCTCCCCGGTGCCTTCACAGATCGGTGGGCCCTTGGAAATTAACACC	343
QY	63	GCGAGCCCCGAGGGCCAGAGAGGCGAGCTGCCGAGCTCTCCGGGGGTCCGCCCG	122
Db	344	GCGAGCCCCCGGGGCCAGAGAGGCGAGCTGCCGAGCTCTCCGGGGGTCCGCCCG	403
QY	123	CGACCTTCTTCTGCGCTTCGCATCTCCCTCCGGGGGTCTTGGACATGCCAGGAATAA	182
Db	404	CGACCTTCTTCTGCGCTTCGCATCTCCCTCCGGGGGTCTTGGACATGCCAGGAATAA	463
QY	183	AAGGATACACTGTACATTCCTGCGCTCTGCTTCCAAAGCCCTGGGAATGCACAGC	242
Db	464	AAGGATACACTGTACATTCCTGCGCTCTGCTTCCAAAGCCCTGGGAATGCACAGC	523
QY	243	ACAGTGCAGGAATGGCTTTCGACGCGGATGCCGACATGAGTGTATTTAGATATGATGA	302
Db	524	ACAGTGCAGGAATGGCTTTCGACGCGGATGCCGACATGAGTGTATTTAGATATGATGA	583
QY	303	ATGCGGAACATCCCGCGAGGCTCCCGAGAGACATGATGTGTATTAACAAATATGCGG	362
Db	584	ATGCGGAACATCCCGCGAGGCTCCCGAGAGACATGATGTGTATTAACAAATATGCGG	643
QY	363	GTATTTATGCAATTCCCGGACAAACCTGTGTATTCGAGGGGCTTACTGAAACCCCTACTC	422
Db	644	GTATTTATGCAATCCCGGACAAACCTGTGTATTCGAGGGGCTTACTGAAACCCCTACTC	703
QY	423	GACCCCTTACACAGTCCGTACCCAGACAGCTGCCCCACACTCTGAGCTCCAAACTATCC	482
Db	704	GACCCCTTACACAGTCCGTACCCAGACAGCTGCCCCACACTCTGAGCTCCAAACTATCC	763
QY	483	CAGCATTCACAGGCTCTTATATGCGCTTTGGATACAGATGATGAAGCAACCAATG	542
Db	764	CAGCATTCACAGGCTCTTATATGCGCTTTGGATACAGATGATGAAGCAACCAATG	823
QY	543	TGTGATGTGGAGAGTGTGCACAGATTTCCACAGTGCACACCCACCCAGATTCGAT	602
Db	824	TGTGATGTGGAGAGTGTGCACAGATTTCCACAGTGCACACCCACCCAGATTCGAT	883
QY	603	CAATACGTGAAGGCGGTACACACTGCTCTGTCACCGACGACGAAATTTGGTTCGGAAGGCA	662
Db	884	CAATACGTGAAGGCGGTACACACTGCTCTGTCACCGACGACGAAATTTGGTTCGGAAGGCA	943
QY	663	GTGCTTGACATTTATGAATTCGCTATGTTACTGCCAGACGCTCTGTGGAATGTTCC	722
Db	944	GTGCTTGACATTTATGAATTCGCTATGTTACTGCCAGACGCTCTGTGGAATGTTCC	1003
QY	723	TGATTCCTATTCCTGTACATCAACCCGTTTAACTCATATGATGGATGGAAAGTCTTG	782
Db	1004	TGATTCCTATTCCTGTACATCAACCCGTTTAACTCATATGATGGATGGAAAGTCTTG	1065
QY	783	CCAAGATGTGAAGAGTGTGCCACCGAAMACCCCTGCTGCAAACTCGTCAACACCTA	842
Db	1064	CCAAGATGTGAAGAGTGTGCCACCGAAMACCCCTGCTGCAAACTCGTCAACACCTA	1122
QY	843	CGGCTCTTTCATTCGCGCTGTGACCCAGGATATGAACCTTGGAGAAATGGCGTTCATTTG	902
Db	1124	CGGCTCTTTCATTCGCGCTGTGACCCAGGATATGAACCTTGGAGAAATGGCGTTCATTTG	1183
QY	903	CAGTATATGAGACAGTGCAGGCTCTCTGATTTCTCTGCCAACATGATGATGTGAACCA	962
Db	1184	CAGTATATGAGACAGTGCAGGCTCTCTGATTTCTCTGCCAACATGATGATGTGAACCA	1244
QY	963	GCCCGGACATATCTTGTCTCTGCGCTCCAGGCTACATCTCTGTGATGACACCGAAG	1022

Db	1244	GC	CCGCGC	CA	CA	TA	CT	TG	CT	CT	CC	CT	CC	CT	CA	AG	CT	TA	CT	CT	GT	GG	AT	GA	CA	CA	CG	AA	G	1303					
OY	1023	CT	GC	CA	AG	CA	TA	CA	GA	AT	GT	G	AG	CA	CA	AG	CA	CA	CT	GC	CA	CT	GC	CA	AG	CA	CT	GC	AG	CA	CT	GT	G	1082	
Db	1304	CT	GC	CA	AG	CA	TA	CA	GA	AT	GT	G	AG	CA	CA	AG	CA	CA	CT	GC	CA	CT	GC	CA	AG	CA	CT	GC	AG	CA	CT	GT	G	1363	
OY	1083	CT	CA	AA	TT	T	AC	AA	G	G	G	G	CT	CA	AA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	1142
Db	1364	CT	CA	AA	TT	T	AC	AA	G	G	G	G	CT	CA	AA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	1423
OY	1143	G	AG	GA	T	CA	GT	TA	AA	CC	G	T	GT	AT	GT	CT	CT	G	T	GA	AA	CC	CT	GT	G	CA	GA	CA	CA	CC	G	1202			
Db	1424	G	AG	GA	T	CA	GT	TA	AA	CC	G	T	GT	AT	GT	CT	CT	G	T	GA	AA	CC	CT	GT	G	CA	GA	CA	CA	CC	G	1483			
OY	1203	CT	TT	AC	CA	CT	TT	AC	CG	GG	GA	CA	T	GA	CA	CG	TC	CA	GA	CA	CG	TC	CA	GA	CA	CG	TC	CA	GA	CA	CG	TC	CA	1262	
Db	1484	CT	TT	AC	CA	CT	TT	AC	CG	GG	GA	CA	T	GA	CA	CG	TC	CA	GA	CA	CG	TC	CA	GA	CA	CG	TC	CA	GA	CA	CG	TC	CA	1543	
OY	1263	CT	TC	CA	AA	T	GA	CA	CG	CA	CA	AG	CC	CT	CA	CC	CT	CG	GG	CC	CT	AT	TA	CA	TT	TC	CA	GA	T	CA	AT	CA	1322		
Db	1544	CT	TC	CA	AA	T	GA	CA	CG	CA	CA	AG	CC	CT	CA	CC	CT	CG	GG	CC	CT	AT	TA	CA	TT	TC	CA	GA	T	CA	AT	CA	1603		
OY	1323	T	GG	GA	AT	GA	GG	G	CA	GA	AA	TT	T	CA	T	CA	T	CG	GG	CA	AA	GG	CC	CC	CA	T	CA	GT	GC	CA	CC	CT	GT	1382	
Db	1604	T	GG	GA	AT	GA	GG	G	CA	GA	AA	TT	T	CA	T	CA	T	CG	GG	CA	AA	GG	CC	CC	CA	T	CA	GT	GC	CA	CC	CT	GT	1663	
OY	1383	G	AT	GA	CA	CG	CC	CA	T	CA	AA	G	G	G	CC	CC	GG	AA	NT	CA	CT	GG	CA	CT	GG	CA	AA	AT	GA	T	CA	CT	GT	1442	
Db	1664	G	AT	GA	CA	CG	CC	CA	T	CA	AA	G	G	G	CC	CC	GG	AA	NT	CA	CT	GG	CA	CT	GG	CA	AA	AT	GA	T	CA	CT	GT	1723	
OY	1443	CA	CA	CA	CT	GA	T	CA	CA	CT	T	CA	GA	GC	CA	AG	CT	CC	GT	AT	CC	GA	CT	CG	CA	T	AT	T	GT	GT	GC	CA	1502		
Db	1724	CA	CA	CA	CT	GA	T	CA	CA	CT	T	CA	GA	GC	CA	AG	CT	CC	GT	AT	CC	GA	CT	CG	CA	T	AT	T	GT	GT	GC	CA	1783		
OY	1503	GT	AC	CA	CT	AT	T	GA	CA	CT	CG	GG	CT	GA	GC	CT	CA	CA	CG	CT	CA	T	GG	CA	CG	CA	CA	GA	G	1562					
Db	1784	GT	AC	CA	CT	AT	T	GA	CA	CT	CG	GG	CT	GA	GC	CT	CA	CA	CG	CT	CA	T	GG	CA	CG	CA	CA	GA	G	1843					
OY	1563	C	AG	A	G	A	G	A	G	A	A	A	T	A	C	A	G	A	G	A															

Db 696 CACGATCTCCAGGCGCTTTATATGCGCTTTGGATACCATGATGTAAGAAAGCAACCATG 755
QY 543 TGTGATGTGGAGAGTGTGCAACAGATTCCTCCACAGTGCACACCCCAAGATCTGCAT 602
Db 756 TGTGGATGTGACACAGTGTGCAACAGATTCCTCCACAGTGCACACCCCAAGATCTGCAT 815
QY 603 CAATACTGAAGGGGGGTACACCTGCTCTGACACGACGAGATTTGGCTTTGGAAGGCCA 662
Db 816 CAATACTGAAGGGGGGTACACCTGCTCTGACACGACGAGATTTGGCTTTGGAAGGCCA 875
QY 663 GTGCTTGACATTTATGATGATGCTGCTGATGTTACTGCTCCAGCAGCTCTGTGCAATGTTCC 722
Db 876 GTGCTTGACATTTATGATGATGCTGCTGATGTTACTGCTCCAGCAGCTCTGTGCAATGTTCC 935
QY 723 TGGATCCTATTCTTGTATGATGCAACCCGTTTACCTCATATGAGATGGAAGTCTTG 782
Db 936 TGGATCCTATTCTTGTATGATGCAACCCGTTTACCTCATATGAGATGGAAGTCTTG 995
QY 783 CCAAGATGTGAAGAGTGTGCGACCGAGAACCCCTGCTGCAAACTGCTCAACACCTA 842
Db 996 CCAAGATGTGAAGAGTGTGCGACCGAGAACCCCTGCTGCAAACTGCTCAACACCTA 1055
QY 843 CGGCTCTTTCATGTCGCGCTGTGACCCAGAGATATGAACTTGAGAAATGGCGTTCAATG 902
Db 1056 CGGCTCTTTCATGTCGCGCTGTGACCCAGAGATATGAACTTGAGAAATGGCGTTCAATG 1115
QY 903 CAGTGATGTGAGAGTGTGACGCTTCTGATGTTCTCTGCCAATGATGATGTAACCA 962
Db 1116 CAGTGATGTGAGAGTGTGACGCTTCTGATGTTCTCTGCCAATGATGATGTAACCA 1175
QY 963 GCGCGGACATTAATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1022
Db 1176 GCGCGGACATTAATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1235
QY 1023 CTGCCAAGACATTAAGAAATGTGAGACAGAGAACCAACAGTGCACACCTGCAAGAGCTG 1082
Db 1236 CTGCCAAGACATTAAGAAATGTGAGACAGAGAACCAACAGTGCACACCTGCAAGAGCTG 1295
QY 1083 CTACAAATTTACAAGGGGGCTTCAAAATGATGATGACCCCATCGCTGTGAGAGGCTTATCT 1142
Db 1296 CTACAAATTTACAAGGGGGCTTCAAAATGATGATGACCCCATCGCTGTGAGAGGCTTATCT 1355
QY 1143 GAGGATGATGATTAACCCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1202
Db 1356 GAGGATGATGATTAACCCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1415
QY 1203 CTTTACCATCTTGTACCGGACATGAGCGTGTGACGAGCGTCCCTTCCCGTGACAT 1262
Db 1416 CTTTACCATCTTGTACCGGACATGAGCGTGTGACGAGCGTCCCTTCCCGTGACAT 1475
QY 1263 CTTCCAAATGCAAGCCGACGACCCGCTACCCCTGGGGCTTATTACATTTTCCAGATCAATC 1322
Db 1476 CTTCCAAATGCAAGCCGACGACCCGCTACCCCTGGGGCTTATTACATTTTCCAGATCAATC 1535
QY 1323 TGGGAATGAGGGGACAGAAATTTTACATGCGGCAAAAGGGGCCCATCATGCTCCACCTGCT 1382
Db 1536 TGGGAATGAGGGGACAGAAATTTTACATGCGGCAAAAGGGGCCCATCATGCTCCACCTGCT 1595
QY 1383 GATGACAGCCGCCATCAAAAGGGGCCGGAATTCAGCTGAGCTTGGAAATGATCACTGT 1442
Db 1596 GATGACAGCCGCCATCAAAAGGGGCCGGAATTCAGCTGAGCTTGGAAATGATCACTGT 1655
QY 1443 CAACACTGTATCAACTTCAGAGGACGCTCCGATCCGACTGCGGATATATGCTGCA 1502
Db 1556 CAACACTGTATCAACTTCAGAGGACGCTCCGATCCGACTGCGGATATATGCTGCA 1715
QY 1503 GTACCATTTGTAGGCTGCGGCTGAGAGCTCCGAGGCTGCTCTCATTTGGCACAAGGGA 1562
Db 1716 GTACCATTTGTAGGCTGCGGCTGAGAGCTCCGAGGCTGCTCTCATTTGGCACAAGGGA 1775
QY 1563 CAGAGAAAGAGAGAAATTAACAGAGATGAGAGGACACAGACTTTGGCAATTTCCG 1622
Db 1776 CAGAGAAAGAGAGAAATTAACAGAGATGAGAGGACACAGACTTTGGCAATTTCCG 1835

QY 1623 CTGAAGCTTTCCCGAAGAGTCAAGCCCGACTTCTGACTCTACCTGTACTATTGACA 1682
Db 1836 CTGAAGCTTTCCCGAAGAGTCAAGCCCGACTTCTGACTCTACCTGTACTATTGACA 1895
QY 1683 CCGTGAACCTGCAAGAGCTTGGCCACCCCACTTCTGATGATGACACTTATCAAAAGTAT 1742
Db 1896 CCGTGAACCTGCAAGAGCTTGGCCACCCCACTTCTGATGATGACACTTATCAAAAGTAT 1955
QY 1743 ATCATGCTCCCGATAGAAGATTTGTTGGAAATTTCAAGGCGCTTATGTTATTTCA 1802
Db 1956 ATCATGCTCCCGATAGAAGATTTGTTGGAAATTTCAAGGCGCTTATGTTATTTCA 2015
QY 1803 CTATTTTCAAAAGAAATAGATTAAGTTTGGGGGGCTGACTATATGTTCAAAAGCTGTG 1862
Db 2016 CTATTTTCAAAAGAAATAGATTAAGTTTGGGGGGCTGACTATATGTTCAAAAGCTGTG 2075
QY 1863 AACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1922
Db 2076 AACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2135
QY 1923 CAAAGACCCGGAGCTGGCGGGGAACCTGGGAGTACTAGTTGCTTTTGGCTGACA 1982
Db 2136 CAAAGACCCGGAGCTGGCGGGGAACCTGGGAGTACTAGTTGCTTTTGGCTGACA 2195
QY 1983 GAGAGGCTATGTAACAAACCAACAGCAGATGAGAGGTTTGAAGATGTGTTCAA 2042
Db 2196 GAGAGGCTATGTAACAAACCAACAGCAGATGAGAGGTTTGAAGATGTGTTCAA 2255
QY 2043 AACCATGCTGTGATTTTCAACCATTAAGAAAGTTAGTTGCTTAAATTTGTATAC 2102
Db 2256 AACCATGCTGTGATTTTCAACCATTAAGAAAGTTAGTTGCTTAAATTTGTATAC 2315
QY 2103 GGTTAATTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2162
Db 2316 GGTTAATTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2375
QY 2163 GAAAGGCTTCAAGACATGCTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2222
Db 2376 GAAAGGCTTCAAGACATGCTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2435
QY 2223 ACCCAGTGTGCTTCTTGAAGGACCCCTTAATCTGCTTCTTGAATTTTACCAGAT 2282
Db 2436 ACCCAGTGTGCTTCTTGAAGGACCCCTTAATCTGCTTCTTGAATTTTACCAGAT 2495
QY 2283 GGATTGGAATGCAAGAGCTTCCAACTGATTAATATTGAAGA 2328
Db 2496 GGATTGGAATGCAAGAGCTTCCAACTGATTAATATTGAAGA 2541

RESULT 4
AX376340
LOCUS AX376340 2609 bp DNA linear PAT 01-MAR-2002
DEFINITION Sequence 407 from Patent WO168848.
ACCESSION AX376340
VERSION AX376340.1 GI:19170561
KEYWORDS
SOURCE
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Baker, K.P., Chen, J., Desnoyers, L., Goddard, A., Godowski, P.J.,
Gurney, A.L., Pan, J., Smith, V., Watanabe, C.K., Wood, W.I. and
Zhang, Z.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
the same
Patent: WO 0168848-A 407 20-SEP-2001;
Genentech, Inc. (US)
JOURNAL Location/Qualifiers
FEATURES
source
1..2609
/organism="Homo sapiens"
/db_xref="taxon:9606"

Db	1236	CTGCCAAGACATCAAGATGTGAGCAGAGAACACGCTGCAACCTGACAGACAGCG	1295
Qy	1083	CTACAAATTACAAAGGGGGCTTCAAAATGCATGACCCCATCCGCTGTGAGAGCCCTTATCT	1142
Db	1296	CTACAAATTACAAAGGGGGCTTCAAAATGCATGACCCCATCCGCTGTGAGAGCCCTTATCT	1355
Qy	1143	GAGATCAGTGAATACCGCTGTATGTCTCTGAGAAACCTGGCTCAGAGACAGCC	1202
Db	1356	GAGATCAGTGAATACCGCTGTATGTCTCTGAGAAACCTGGCTCAGAGACAGCC	1415
Qy	1203	CTTTACCATCTTGTACCGGGACATGACGTGTGTGAGAGAGCTCCGTTCCCGCTGACAT	1262
Db	1416	CTTTACCATCTTGTACCGGGACATGACGTGTGTGAGAGAGCTCCGTTCCCGCTGACAT	1475
Qy	1263	CTTCCAAATGCAACGACGACCCGCTACCCCTGGGGCCCTATTACATTTTCCAGATCAAAATC	1322
Db	1476	CTTCCAAATGCAACGACGACCCGCTACCCCTGGGGCCCTATTACATTTTCCAGATCAAAATC	1535
Qy	1323	TGGGAATGAGGGGAGAGAAATTTTACATGCGCAAAAGGGCCCATCAGTGCACCCCTGT	1382
Db	1536	TGGGAATGAGGGGAGAGAAATTTTACATGCGCAAAAGGGCCCATCAGTGCACCCCTGT	1595
Qy	1383	GATGACACGCCCCATCAAAAGGGCCCGGGAATCCAGCTGACCTTGAATGATCAGTGT	1442
Db	1596	GATGACACGCCCCATCAAAAGGGCCCGGGAATCCAGCTGACCTTGAATGATCAGTGT	1655
Qy	1443	CACACTGTCAATCACTTCAGAGCAGCTCCGCTGATCCGATGCGGAGATATGTGTCCGA	1502
Db	1656	CACACTGTCAATCACTTCAGAGCAGCTCCGCTGATCCGATGCGGAGATATGTGTCCGA	1715
Qy	1503	GTACCCATCTGAGCTCGGGCTGAGACCTCCGACGCTGCTCATATGGCCACCAAGGA	1562
Db	1716	GTACCCATCTGAGCTCGGGCTGAGACCTCCGACGCTGCTCATATGGCCACCAAGGA	1775
Qy	1563	CAGGAGAGAGAGAAATTAACAGAGAAATGAGAGCGACACAGACGTTAGGCATTTCTGT	1622
Db	1776	CAGGAGAGAGAGAAATTAACAGAGAAATGAGAGCGACACAGACGTTAGGCATTTCTGT	1835
Qy	1623	CTGAACGTTTCCCGGAAGAGTCAAGCCCGGACCTCTGCTCACTCACTTACTATTTGCGA	1682
Db	1836	CTGAACGTTTCCCGGAAGAGTCAAGCCCGGACCTCTGCTCACTCACTTACTATTTGCGA	1895
Qy	1683	CCGTGACCCCTGAGAGACTTGCACACCCCAAGTCTCTATGATACAGTATCAAAAAAGTAT	1742
Db	1896	CCGTGACCCCTGAGAGACTTGCACACCCCAAGTCTCTATGATACAGTATCAAAAAAGTAT	1955
Qy	1743	ATCATTTGCTCCCTGATAGAGATTTGTTGTAATTTCAAGCCCTCAGTTTATTTCCA	1802
Db	1956	ATCATTTGCTCCCTGATAGAGATTTGTTGTAATTTCAAGCCCTCAGTTTATTTCCA	2015
Qy	1803	CTATTTTCAAGAAATAGATTAAGTTTGGGGGGTCTGAGTATGTTCAAAAGCTGTG	1862
Db	2016	CTATTTTCAAGAAATAGATTAAGTTTGGGGGGTCTGAGTATGTTCAAAAGCTGTG	2075
Qy	1863	AACAGCTGCTGCTCACTTCTTCACTCTTCCACTCTCTCACTGTTACTCTTGTG	1922
Db	2076	AACAGCTGCTGCTCACTTCTTCACTCTTCCACTCTCTCACTGTTACTCTTGTG	2135
Qy	1923	CAAAAGCCCGGAGCTGGCGGGGAACCTTGGAGTACCTTGTGCTTTTGGCTTACCA	1982
Db	2136	CAAAAGCCCGGAGCTGGCGGGGAACCTTGGAGTACCTTGTGCTTTTGGCTTACCA	2195
Qy	1983	GAGAAGCTATGTAACAACAACAGAGAGATGGAAGGGTTTGAAGAATGTGTTCAA	2042
Db	2196	GAGAAGCTATGTAACAACAACAGAGAGATGGAAGGGTTTGAAGAATGTGTTCAA	2255
Qy	2043	AACCATGCTGTATTTTCAACCATATAAAGAGTTTCACTGTCTTAAATTTGTATAC	2102
Db	2256	AACCATGCTGTATTTTCAACCATATAAAGAGTTTCACTGTCTTAAATTTGTATAC	2315
Qy	2103	GGTTTATTTCTGCTTGTTCATTTTGAATTTTAAATAATATGTGTAATTCCTTC	2162
Db	2316	GGTTTATTTCTGCTTGTTCATTTTGAATTTTAAATAATATGTGTAATTCCTTC	2375
Qy	2163	GAAAGGCTTCAGACACATGCTATGTCTGTCTTCCCAAAACCATGCTCTCCATTTT	2222
Db	2376	GAAAGGCTTCAGACACATGCTATGTCTGTCTTCCCAAAACCATGCTCTCCATTTT	2435
Qy	2223	AGCCAGTGTCTTCTTTGAGAGACCCCTTAACCTTCTTCTTTAGAAATTTTACCAGT	2282
Db	2436	AGCCAGTGTCTTCTTTGAGAGACCCCTTAACCTTCTTCTTTAGAAATTTTACCAGT	2495
Qy	2283	GGATTGGAATGACAGAGCTCTCCAAACGATTAATTAATTTGAAGA	2328
Db	2496	GGATTGGAATGACAGAGCTCTCCAAACGATTAATTAATTTGAAGA	2541
RESULT 3 AR173204 LOCUS AR173204 DEFINITION Sequence 2 from patent US 6303765. ACCESSION AR173204 VERSION AR173204.1 GI:17912695 KEYWORDS SOURCE ORGANISM Unknown. REFERENCE 1 (bases 1 to 2550) Bandman O., Corley N.C. and Guejler K.J. Human extracellular matrix proteins Patent: US 6303765-A 2.16-Oct-2001; JOURNAL FEATURES source 1..2550 /organism="unknown" BASE COUNT 623 a 690 c 596 g 641 t ORIGIN			
Query Match Best Local Similarity 99.8%; Score 2324.4; DB 6; Length 2550; Matches 2325; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	3	CCCGGCGCTCTCCCGGCTCTCTCTCCAGACCTCGTGGCCCTCGAATTAACACCC	62
Db	216	CCCGGCGCTCTCCCGGCTCTCTCTCCAGACCTCGTGGCCCTCGAATTAACACCC	275
Qy	63	GCGAGCCCGGAGGGCCAGAGAGGCGCAGCTGCCGAGCTCTCGGGGGTCCGCCCG	122
Db	276	GCGAGCCCGGAGGGCCAGAGAGGCGCAGCTGCCGAGCTCTCGGGGGTCCGCCCG	335
Qy	123	CGAGCTTCTTCTGCGCTTGCATCTCTCTCGCGCGCTTGTGACATGCCAGGAATAA	182
Db	336	CGAGCTTCTTCTGCGCTTGCATCTCTCTCGCGCGCTTGTGACATGCCAGGAATAA	395
Qy	183	AAGATATCTACCTTACATCTTGGCTCTCTCTTCCAAACCTCGGAATGACAGGC	242
Db	396	AAGATATCTACCTTACATCTTGGCTCTCTCTTCCAAACCTCGGAATGACAGGC	455
Qy	243	ACAGTGCAGAAATGCTTGTGACTGATGCGCAGTCCAGAGACAGTGTATGATATGATGA	302
Db	456	ACAGTGCAGAAATGCTTGTGACTGATGCGCAGTCCAGAGACAGTGTATGATATGATGA	515
Qy	303	ATGCCGAAACATCCCGAGAGCTTCCGAGAGACATGATGTGTAAACCAAAATGGCG	362
Db	516	ATGCCGAAACATCCCGAGAGCTTCCGAGAGACATGATGTGTAAACCAAAATGGCG	575
Qy	363	GTAATTTATGATTCCTCCGGAACAACCTGTATGAGAGGCTCTACTGAAACCCCTACTC	422
Db	576	GTAATTTATGATTCCTCCGGAACAACCTGTATGAGAGGCTCTACTGAAACCCCTACTC	635
Qy	423	GACCCCTACTCAGAGTCCGTACCAGAGAGTGCACCACTCTCAGCTCCAAACTATCC	482
Db	636	GACCCCTACTCAGAGTCCGTACCAGAGAGTGCACCACTCTCAGCTCCAAACTATCC	695
Qy	483	CACGATCTCAGAGCTCTTATATGCGGCTTGTGATACAGATGATGAAGCAACCAATG	542

QY	1561	GACAGGGAAGAGAGGAAATTAACAGAGAAATGAAGAGCAACAGACAGTTAGCATTTCC	1620
Db	1561	GACAGGGAAGAGAGGAAATTAACAGAGAAATGAAGAGCAACAGACAGTTAGCATTTCC	1620
QY	1621	TGCTGAACGTTTCCCGAAGAGTAGTAGCCCGCACTTCCAGATCTCACTCTAGTATTGCA	1680
Db	1621	TGCTGAACGTTTCCCGAAGAGTAGTAGCCCGCACTTCCAGATCTCACTCTAGTATTGCA	1680
QY	1681	GACCTGTACCCCTCAGAGACTTGGCCACCCCAAGTTCATATGATACAGTTATCAAAAAGTA	1740
Db	1681	GACCTGTACCCCTCAGAGACTTGGCCACCCCAAGTTCATATGATACAGTTATCAAAAAGTA	1740
QY	1741	TTATTCATTCGTCCTCCCTGATTAATAAGTTTGTGTGAATTTTCAAGGCTTCAGTTATTTC	1800
Db	1741	TTATTCATTCGTCCTCCCTGATTAATAAGTTTGTGTGAATTTTCAAGGCTTCAGTTATTTC	1800
QY	1801	CACATATTTTCAAGAAAATAGATTAGTTTGGGGGGGTGTGAGTCTATGTTCAAAGCTG	1860
Db	1801	CACATATTTTCAAGAAAATAGATTAGTTTGGGGGGGTGTGAGTCTATGTTCAAAGCTG	1860
QY	1861	TGAACAGCTTGTCTCACTTCTTCACTCTTCCACTCTCTCTCACTGTATTAGTCTT	1920
Db	1861	TGAACAGCTTGTCTCACTTCTTCACTCTTCCACTCTCTCTCACTGTATTAGTCTT	1920
QY	1921	TGCAAAAGACCCGGAGAGCTGGGGGGGAACCCCTGGAGTAGCTAGTTTGCTTTTGGCGTACA	1980
Db	1921	TGCAAAAGACCCGGAGAGCTGGGGGGGAACCCCTGGAGTAGCTAGTTTGCTTTTGGCGTACA	1980
QY	1981	CAGAGAAGGCTATGTAAACAACACACAGAGATGAAGGGTTTTAGAGATGTGTTC	2040
Db	1981	CAGAGAAGGCTATGTAAACAACACACAGAGATGAAGGGTTTTAGAGATGTGTTC	2040
QY	2041	AAACCATGCTGTATTTTCAACCATTAAGAAAGTTTCAGTTGCTTAAATTTGTATA	2100
Db	2041	AAACCATGCTGTATTTTCAACCATTAAGAAAGTTTCAGTTGCTTAAATTTGTATA	2100
QY	2101	ACGGTTTAATTCGTCTGTCTCATATTGAGATTTTAAAAAATAGTGTGATAATTTGCTT	2160
Db	2101	ACGGTTTAATTCGTCTGTCTCATATTGAGATTTTAAAAAATAGTGTGATAATTTGCTT	2160
QY	2161	TGCAAAAGGCTTCAGACACATGCTATGTCTGTCTCCCAAAACCCAGTCTCCGTCAT	2220
Db	2161	TGCAAAAGGCTTCAGACACATGCTATGTCTGTCTCCCAAAACCCAGTCTCCGTCAT	2220
QY	2221	TTAGCCAGTGTCTTTCTTTGAGAGACCCCTTAATCTTGTCTTTTGAATTTTACCAA	2280
Db	2221	TTAGCCAGTGTCTTTCTTTGAGAGACCCCTTAATCTTGTCTTTTGAATTTTACCAA	2280
Y	2281	TTGGATTGGAAATCAGAGAGTCTCCAAACTGATTAATTTGAAGAAGA 2328	
Db	2281	TTGGATTGGAAATCAGAGAGTCTCCAAACTGATTAATTTGAAGAAGA 2328	
RESULT 2			
AR036548			
LOCUS	AR036548	2550 bp	DNA
DEFINITION	Sequence 2 from patent US 5872234.		
ACCESSION	AR036548		
VERSION	AR036548.1	GI:5953216	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 2550)		
AUTHORS	Bandman,O., Corley,N.C. and Guegler,K.J.		
TITLE	Human extracellular matrix proteins		
JOURNAL	Patent: US 5872234-A 2 16-FEB-1999;		
FEATURES	Location/Qualifiers		
source	1..2550		
BASE COUNT	623 a	690 c	596 g 641 t
ORIGIN	/organism="unknown"		

Query Match	99.8%	Score 2324.4	DB 6	Length 2350	
Best Local Similarity	100.0%	Pred. No. 0			
Matches 2352	Conservative 0	Mismatches 1	Indels 0	Gaps 0	
QY	3	CCCCGGCGTCCCGCGTCCCTCCACAGACTCGCTCGGCCCTCTGGAAATTAACACCC	62		
DB	216	CCCCGGCGCTCCCGCGTCCCTCCACAGACTCGCTCGGCCCTCTGGAAATTAACACCC	275		
QY	63	GGAGACCCCGGAGGGGCCAGAGAGAGCGACAGTCCCGGAGCTCTCGGGGGTCCCGCCG	122		
DB	276	GGAGACCCCGGAGGGGCCAGAGAGAGCGACAGTCCCGGAGCTCTCGGGGGTCCCGCCG	335		
QY	123	CGAGCTTCTCTCGCTTCGCTTCCTCTCGCGGCTCTTGGACATGCCAGAAATAA	182		
DB	336	CGAGCTTCTCTCGCTTCGCTTCCTCTCGCGGCTCTTGGACATGCCAGAAATAA	395		
QY	183	AAGGATACACTGTATTACATTCCTGGCTCTGCTCTTCCAGACCTGGGAATGCACAGC	242		
DB	336	AAGGATACACTGTATTACATTCGCTCTGCTCTTCCAGACCTGGGAATGCACAGC	455		
QY	243	ACAGTGCAGATGGCTTGGACCTGGACAGTCCGACAGCAGACAGTGTATATGATGA	302		
DB	456	ACAGTGCAGATGGCTTGGACCTGGACAGTCCGACAGCAGACAGTGTATATGATGA	515		
QY	303	ATGCCGAACATCCCCGAGCCTCGCGAGAGACATGATGTGTAAACAAATATGCGG	362		
DB	516	ATGCCGAACATCCCCGAGCCTCGCGAGAGACATGATGTGTAAACAAATATGCGG	575		
QY	363	GTAATTATGACATCCCGGAGAACCCGTATGTAGAGGGCTTACTGAAACCTTACTC	422		
DB	576	GTAATTATGACATCCCGGAGAACCCGTATGTAGAGGGCTTACTGAAACCTTACTC	635		
QY	423	GACCCCTTACACAGGTCCGCTACCCAGCAGCTGCCCCACACTTCAGCTCCAAACTATCC	482		
DB	636	GACCCCTTACACAGGTCCGCTACCCAGCAGCTGCCCCACACTTCAGCTCCAAACTATCC	695		
QY	483	CACGATTCACAGGCTCTTATATGCGCTTTGGATACCATGTGATGAAGCAACCAATG	542		
DB	696	CACGATTCACAGGCTCTTATATGCGCTTTGGATACCATGTGATGAAGCAACCAATG	755		
QY	543	TGTGGATGTGGAGGTGGCAACGATTTCCACACAGTGCACACCCACCGAGATCTGAT	602		
DB	756	TGTGGATGTGGAGGTGGCAACGATTTCCACACAGTGCACACCCACCGAGATCTGAT	815		
QY	603	CAATACTGAGAGCGGGTACACCTGCTCTGACACGACGAGATTTGGCTTCGGAAGGCA	662		
DB	816	CAATACTGAGAGCGGGTACACCTGCTCTGACACGACGAGATTTGGCTTCGGAAGGCA	875		
QY	663	GTGCTTACATGTGATGATGTGCTATGTGTTACTGCCAGCAGCTCTGTGCAATGTTC	722		
DB	876	GTGCTTACATGTGATGATGTGCTATGTGTTACTGCCAGCAGCTCTGTGCAATGTTC	935		
QY	723	TGGAATCCATCTGTATACATCAACCCGTGTTAACCTCAATAGATGGAAGTCTTG	782		
DB	936	TGGAATCCATCTGTATACATCAACCCGTGTTAACCTCAATAGATGGAAGTCTTG	995		
QY	783	CCAAGATGTGACAGTGTGCGACACGAGAACCCCTGCTGCAAACTCTGCAACACCTA	842		
DB	996	CCAAGATGTGACAGTGTGCGACACGAGAACCCCTGCTGCAAACTCTGCAACACCTA	1055		
QY	843	CGGCTCTTTCATCTGCGCGTGTGACCCAGATATGAACTTGAAGAAATGGCGTTCATG	902		
DB	1056	CGGCTCTTTCATCTGCGCGTGTGACCCAGATATGAACTTGAAGAAATGGCGTTCATG	1115		
QY	903	CAGGATATGAGACAGTGCACCTCTCGATTCCTCTGCCAATATGATGTGTGAACA	962		
DB	1116	CAGGATATGAGACAGTGCACCTCTCGATTCCTCTGCCAATATGATGTGTGAACA	1175		
QY	963	GCCTGGGACATATCTTGTGCTCTGCTCCGCTCCAGGCTATCTCTGTGTGATCAACCGAG	1022		
DB	1176	GCCTGGGACATATCTTGTGCTCTGCTCCGCTCCAGGCTATCTCTGTGTGATCAACCGAG	1235		
QY	1023	CTGGCAAGACATCAAGATGTGAGCCAGGAAACACAGTGCACACCTGACAGACGTG	1082		

Matsumori, A., Sasayama, S., Chien, K.R. and Honjo, T.
DANCE, a novel secreted RGD protein expressed in developing,
atherosclerotic, and balloon-injured arteries
J. Biol. Chem. 274 (32), 22476-22483 (1999)
MEDLINE 99357779
PUBMED 10428823
2 (bases 1 to 2328)
REFERENCE Nakamura, T., Yabe, D., Tashiro, K. and Honjo, T.
AUTHORS Submitted (09-DEC-1998) Medicine, University of California San
TITLE Direct Submission Diego, 9500 Gilman Drive, La Jolla, CA 92093-0613, USA
JOURNAL location/Qualifiers
1. 2328
FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="14"
/map="14q32.1"
169..1515
/note="DANCE: contains 6 cbpGF domains and 1 RGD motif;
secreted protein; expressed in developmental aorta, neural
crest cells, balloon injured vessels, atherosclerotic
lesions"
/codon_start=-1
/product="developmental arteries and neural crest EGF-like
protein"
/protein_id="AA041768.1"
/db_xref="GI:5305673"
/translation="MPGIRKRLITVITLALCLPSPRNOAOCTNGEDLRSGCDDID
ECPTRPACQDMVQNGYLCIPRNPYRGPSNPSTSPYSPAPPAAPLSAP
NPTISPLRFGYOMDSQVDEACSHQCNPTQIDINTEGVCSTQDYM
LLEGOCIDIDECRGYCOQLCANPAGSYSCNPNFTLNEGSCODVEACENPCV
QTCVNTYGFICRCDPGELEDEGHCSDMECSFEFLCQHECVNOPTIYCSPPG
YLLDNRSCODINECHRNHTCNLQOCVYVLOGGFKCIDPRCEPPIYISDNRCM
PANNPCGRDPPFLYRDMDVYSSRSVPADIFQMAITRYREPAITIFQISNEGHEF
YMQTGISLIVMTPLIKGPRQLDLEMTIVNINFRSSVIRLIVSQPR"

BASE COUNT 560 a 646 c 541 g 581 t

ORIGIN

Query Match 100.0%; Score 2328; DB 9; Length 2328;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCCGGCGCTCCCGCTGCTCTCCACAGACTGCTCGGCCCTCTGGAATAAAACAC 60.
DB 1 GACCCGGCGCTCCCGCTGCTCTCCACAGACTGCTCGGCCCTCTGGAATAAAACAC 60
QY 61 CCGCGAG 120
DB 61 CCGCGAG 120
QY 121 CCGGAGCTTCTCTGCGCTTCCGATCTCTCTGCGCGCTTGGACATGCGCAGAGATA 180
DB 121 CCGGAGCTTCTCTGCGCTTCCGATCTCTCTGCGCGCTTGGACATGCGCAGAGATA 180
QY 181 AAAAGACTACTGTTACATCTGCGCTCTGCTTCCAAAGCCCTGGGAATGACAG 240
DB 181 AAAAGACTACTGTTACATCTGCGCTCTGCTTCCAAAGCCCTGGGAATGACAG 240
QY 241 GCACAGAGCAGATGGCTTGGACCTGATGAGCCAGTACAGACAGTGTATATATGAT 300
DB 241 GCACAGAGCAGATGGCTTGGACCTGATGAGCCAGTACAGACAGTGTATATATGAT 300
QY 301 GAATGCGCAACATCCCGAGAGCTGCGGAGAGACATGATGTGTAAACAAATGGC 360
DB 301 GAATGCGCAACATCCCGAGAGCTGCGGAGAGACATGATGTGTAAACAAATGGC 360
QY 361 GGGTATTATGATTCCTCCGAGACAAACCTGTATGAGAGGGCCCTACTGAAACCCCTAC 420
DB 361 GGGTATTATGATTCCTCCGAGACAAACCTGTATGAGAGGGCCCTACTGAAACCCCTAC 420
QY 421 TCGAAGCCCTACTACAGTCCGTACCCACAGACAGTCCGACCACTCTCAAACTAT 480
DB 421 TCGAAGCCCTACTACAGTCCGTACCCACAGACAGTCCGACCACTCTCAAACTAT 480

QY 481 CCCAGATCTCCAGAGCCCTTATATAGCCGCTTGGATACCAATGGATGAAGAACACAA 540
DB 481 CCCAGATCTCCAGAGCCCTTATATAGCCGCTTGGATACCAATGGATGAAGAACACAA 540
QY 541 TGTGTGATGTGAGAGAGTGTGCAACAGATTCCACAGATGAGACCCACCAAGATCTGC 600
DB 541 TGTGTGATGTGAGAGAGTGTGCAACAGATTCCACAGATGAGACCCACCAAGATCTGC 600
QY 601 ATCAATACGTAAGCGGGGTACACCTGCTCCACGACGAGATATTTGCTTGGAAAGC 660
DB 601 ATCAATACGTAAGCGGGGTACACCTGCTCCACGACGAGATATTTGCTTGGAAAGC 660
QY 661 CAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 661 CAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 721 CCTGATCTCTATTTCTTGTACATGCAACCCCTGTTTACCTCAATGAGATGAGAGTCT 780
DB 721 CCTGATCTCTATTTCTTGTACATGCAACCCCTGTTTACCTCAATGAGATGAGAGTCT 780
QY 781 TGCCAAAGATGGAAGAGAGTGTGCCACGAGAAACCCCTGCTGCAAACTGCTCAACACC 840
DB 781 TGCCAAAGATGGAAGAGAGTGTGCCACGAGAAACCCCTGCTGCAAACTGCTCAACACC 840
QY 841 TAGGGCTCTTCTCATCTGCGCGCTGTGACCCAGATATGAACTTGAGAGAGATGGCTTCAT 900
DB 841 TAGGGCTCTTCTCATCTGCGCGCTGTGACCCAGATATGAACTTGAGAGAGATGGCTTCAT 900
QY 901 TGCAGATATATGAGAGAGTGTGACGCTTCTGTGATTCCTGTGCAACATGATGTGTGAA 960
DB 901 TGCAGATATATGAGAGAGTGTGACGCTTCTGTGATTCCTGTGCAACATGATGTGTGAA 960
QY 961 CAGCCCGGACATATCTTCTGCTCCGCTCCAGCTATATCTGCTGATGATGACACCA 1020
DB 961 CAGCCCGGACATATCTTCTGCTCCGCTCCAGCTATATCTGCTGATGATGACACCA 1020
QY 1021 AGCTGCCAAGACATCAACGAATGTGAGACAGAACCAACAGCTGCAACCTGACAGACAG 1080
DB 1021 AGCTGCCAAGACATCAACGAATGTGAGACAGAACCAACAGCTGCAACCTGACAGACAG 1080
QY 1081 TGCTACAAATTTACAAAGGGGCTTCAAAATGATGACACCCATCGCTGTAGAGAGCTTAT 1140
DB 1081 TGCTACAAATTTACAAAGGGGCTTCAAAATGATGACACCCATCGCTGTAGAGAGCTTAT 1140
QY 1141 CTGAGATACATGATTAACCGCTGTATGTGCTGTGAGAACCTGCGTGCAGAGACAG 1200
DB 1141 CTGAGATACATGATTAACCGCTGTATGTGCTGTGAGAACCTGCGTGCAGAGACAG 1200
QY 1201 CCTTTTACATCTTGTACCGGAGACATGAGACGTGTGTCAGACGCTCCGCTCCGCTGAC 1260
DB 1201 CCTTTTACATCTTGTACCGGAGACATGAGACGTGTGTCAGACGCTCCGCTCCGCTGAC 1260
QY 1261 ATCTTCCAAATGCAAGCCAGACAGACCCGCTACCTGCGGCTATTAATCTTCCAGATCAA 1320
DB 1261 ATCTTCCAAATGCAAGCCAGACAGACCCGCTACCTGCGGCTATTAATCTTCCAGATCAA 1320
QY 1321 TCTGGAATGAGGGGAGAGAAATTTACATGGCGCAAGGGGCCCTCAGTGCACCCCTG 1380
DB 1321 TCTGGAATGAGGGGAGAGAAATTTACATGGCGCAAGGGGCCCTCAGTGCACCCCTG 1380
QY 1381 GTGATGACACGCCCATCAAAAGGGCCCGGGAATTCACACTGACTTGGAAATGATCACT 1440
DB 1381 GTGATGACACGCCCATCAAAAGGGCCCGGGAATTCACACTGACTTGGAAATGATCACT 1440
QY 1441 GTCAACACTGTATCACTTCAAGAGGACACTCCGATCCGATCCGATGATATATGTGTG 1500
DB 1441 GTCAACACTGTATCACTTCAAGAGGACACTCCGATCCGATCCGATGATATATGTGTG 1500
QY 1501 CAGTACCAATCTGAGGCTCGGGGCTGAGGCTCCGACGCTGCTCATTTGGACCAAGG 1560
DB 1501 CAGTACCAATCTGAGGCTCGGGGCTGAGGCTCCGACGCTGCTCATTTGGACCAAGG 1560

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 08:43:26 ; Search time 6070.42 Seconds

(Without alignments)
11160.894 Million cell updates/sec

Title: US-09-674-379A-12

Perfect score: 2328
Sequence: 1 gaccggcgctctcccgctg.....tgattaatattgaagaga 2328

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Geneml: 1: gb_ba: 2: gb_hlg: 3: gb_in: 4: gb_om: 5: gb_ov: 6: gb_pat: 7: gb_ph: 8: gb_pl: 9: gb_pr: 10: gb_ro: 11: gb_sts: 12: gb_sy: 13: gb_un: 14: gb_vl: 15: em_ba: 16: em_fun: 17: em_hum: 18: em_in: 19: em_mu: 20: em_om: 21: em_or: 22: em_ov: 23: em_pat: 24: em_ph: 25: em_pl: 26: em_ro: 27: em_sts: 28: em_un: 29: em_vl: 30: em_hlg_hum: 31: em_hlg_inv: 32: em_hlg_other: 33: em_hlg_mus: 34: em_hlg_pin: 35: em_hlg_rtd: 36: em_hlg_mam: 37: em_hlg_vrt: 38: em_sy: 39: em_hgo_hum: 40: em_hgo_mus: 41: em_hgo_other:

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2328	100.0	2328	9 AF112152	AF112152 Homo sapi
2	2324.4	99.8	2550	6 AR036548	AR036548 Sequence
3	2324.4	99.8	2550	6 AR173204	AR173204 Sequence
4	2322.8	99.8	2609	6 AX376340	AX376340 Sequence
5	2322.8	99.8	2609	6 AX403659	AX403659 Sequence
6	2321.2	99.7	2367	9 BC022280	BC022280 Homo sapi
7	2014.8	86.5	2019	9 HSA13490	HSA13490 Homo sapi
8	1955.6	84.0	2019	9 AF093118	AF093118 Homo sapi
9	1674.2	71.9	1720	6 A84086	A84086 Sequence 1
10	1481.8	63.7	2230	10 BC006636	BC006636 Mus muscu
11	1481.8	63.7	2478	10 AF112151	AF112151 Mus muscu
12	1476.6	63.4	2499	10 AF112153	AF112153 Rattus no
13	1475.2	63.4	2304	10 AF137350	AF137350 Rattus no
14	974.4	41.9	171901	9 AC007057	AC007057 Homo sapi
15	974.4	41.9	175210	9 CNS07EG5	AL590338 Human chr
16	973.4	41.8	2324	9 AK094032	AK094032 Homo sapi
17	630.4	27.1	648	9 HSM800793	AL110148 Homo sapi
18	501	21.5	172341	2 AC095490	AC095490 Rattus no
19	387.6	16.6	1440	9 HSA132819	AJ132819 Homo sapi
20	387.6	16.6	1536	9 AB030655	AB030655 Homo sapi
21	386	16.6	1480	6 AX023967	AX023967 Sequence
22	386	16.6	1561	9 BC010456	BC010456 Homo sapi
23	386	16.6	1707	9 AK000980	AK000980 Homo sapi
24	386	16.6	1875	6 AX201326	AX201326 Sequence
25	386	16.6	1875	6 AX464184	AX464184 Sequence
26	386	16.6	2018	6 AR095382	AR095382 Sequence
27	384.4	16.5	1757	9 AF109121	AF109121 Homo sapi
28	379.6	16.3	1825	9 AF124486	AF124486 Homo sapi
29	378.6	16.3	1705	9 AF093119	AF093119 Homo sapi
30	353.4	15.2	1778	10 AF046870	AF046870 Cricethlu
31	352.4	15.1	1358	6 AX023976	AX023976 Sequence
32	350.2	15.0	1512	10 AF104223	AF104223 Mus muscu
33	350.2	15.0	1513	6 AX023961	AX023961 Sequence
34	350.2	15.0	1794	10 AF109122	AF109122 Mus muscu
35	348.6	15.0	1724	10 BC012269	BC012269 Mus muscu
36	348.2	15.0	518	6 AX079354	AX079354 Sequence
37	321.8	13.8	340	6 AX332820	AX332820 Sequence
38	321.8	13.8	340	6 AX334715	AX334715 Sequence
39	321.8	13.8	340	6 AX336178	AX336178 Sequence
40	315	13.5	1021	6 AX023954	AX023954 Sequence
41	295.8	12.7	2613	9 AK098186	AK098186 Homo sapi
42	268	11.5	1132	6 AX023965	AX023965 Sequence
43	260.6	11.2	172674	9 CDS0000R	AL049872 Human chr
44	251	10.8	1677	10 BC031184	BC031184 Mus muscu
45	250.6	10.8	1911	10 D89730	D89730 Rattus ratt

ALIGNMENTS

RESULT 1
AF112152
LOCUS
DEFINITION Homo sapiens developmental arteries and neural crest EGF-like protein mRNA, complete cds.
ACCESSION AF112152
VERSION AF112152
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 2328)
Nakamura,T., Ruiz-Lozano,P., Lindner,V., Yabe,D., Tanikawa,M., Furukawa,Y., Koduke,K., Tashiro,K., Lu,Z., Andon,N.L., Schaub,R.,

Not1; Site 2; Sal1; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TGCAGCCAGCGTCCG-3' and
5'-GACTTATCTTCAATTCGGAGCGGCCGCTT-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."

BASE COUNT 225 a 284 c 245 g 179 t
ORIGIN

Query Match 42.9%; Score 576.4; DB 14; Length 933;
Best Local Similarity 94.6%; Pred. No. 9e-149;
Matches 619; Conservative 0; Mismatches 31; Indels 4; Gaps 2;

```
QY 602 ATGGAAGCTCTTGCCCAAGATGTAGACGAGTGTGCGACCGAGAACCCCTGCGCAACCT 661
DB 2 ATGGAAGCTCTTGCCCAAGATGTAGACGAGTGTGCGACCGAGAACCCCTGCGCAACCT 61
662 GCGTCAACACCTACCGCTCTTTCATCTGCGCGTGTGACCCCGAGATGAACTTGGGAG 721
DB 62 GCGTCAACACCTACCGCTCTTTCATCTGCGCGTGTGACCCCGAGATGAACTTGGGAG 121
QY 722 ATGGCGTTCATTGCAATGATATGAGAGAGTGCAGCTTCTGAGTTCCTTGCCAAATG 781
DB 122 ATGGCGTTCATTGCAATGATATGAGAGAGTGCAGCTTCTGAGTTCCTTGCCAAATG 181
QY 782 AGTGTGTGAACCAAGCCCGGCACATCTTGTGCTCTGCTGCTCCCTCCAGGCTACATCCTGTGG 841
DB 182 AGTGTGTGAACCAAGCCCGGCACATCTTGTGCTCTGCTGCTCCCTCCAGGCTACATCCTGTGG 241
QY 842 ATGACAAACGAGCTGCAAGATCAAGAAATGTGAGACAGAGAACCAACAGTGCACAC 901
DB 242 ATGACAAACGAGCTGCAAGATCAAGAAATGTGAGACAGAGAACCAACAGTGCACAC 301
QY 902 TGCAGAGAGAGTGTCAATTTTACAAGGGGGCTTCAATGCATGACCCCAATCCGCTGTG 961
DB 302 TGCAGAGAGAGTGTCAATTTTACAAGGGGGCTTCAATGCATGACCCCAATCCGCTGTG 361
QY 962 AGGAGCCTTATCTGAGAGATCAGTATACCGCTGTATGTCTGCTGAGAAACCTGTGGCT 1021
DB 362 AGGAGCCTTATCTGAGAGATCAGTATACCGCTGTATGTCTGCTGAGAAACCTGTGGCT 421
QY 1022 GCAGAGACAGCCCTTACCATCTTGTACGGGGACATGAGACGTGTGCAGAGAGCTCCG 1081
DB 422 GCAGAGACAGCCCTTACCATCTTGTACGGGGACATGAGAGGTGTGCAGAGAGCTCCG 481
QY 1082 TTCCCGCTGACATCTTCAAAATGCAAGCCAGACCCGCTACCTGCGGCTATTACATTT 1141
DB 482 TTCCCGCTGACATCTTCAAAATGCAAGCCAGACCCGCTACCTGCGGCTATTACATTT 541
QY 1142 TCCAGATCAAAATCTGGGAATGA-GGGCAGAGAAATTTTACATGGCGCAAA---CGGGCCCC 1197
DB 542 TCCAGATCAAAATCTGGGAATGAAGAGGAGCAAGAAATTTTACATGGCGCAAAACGGGGCCCC 601
QY 1198 ATCAGTCCAGCCCTGTGATGACACGCCCATCAAGAGGCCCGCGGAATCCAG 1251
DB 602 ATCAGTCCAGCCCGGGTTGATGAGACCCGCCCATTTCCAAAGGGGGCCCCGG 655
```

Search completed: July 3, 2003, 17:36:03
Job time : 1988.41 secs

NotI; Site-2; SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCAGCGTCG-3' and
5'-GACTGATTCAGATCGGAGCGGCCCTT-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."

BASE COUNT 181 a 266 c 206 g 182 t

ORIGIN

Query Match 46.1%; Score 619; DB 14; Length 835;

Best Local Similarity 98.8%; Pred. No. 1.3e-160;

Matches 645; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

1 ATGCCAGGATTAAGGATGACCTGATACCATTTGCTCTCTCTTCCAGCCCT 60
|||||
171 ATGCCAGGATTAAGGATGACCTGATACCATTTGCTCTCTCTTCCAGCCCT 230
61 GGGATGACAGGACAGTGCAGCAATGGCTTGCATGCGCAGTCAAGACAGTGT 120
|||||
231 GGGATGACAGGACAGTGCAGCAATGGCTTGCATGCGCAGTCAAGACAGTGT 290
121 TTGATATGATGATATGCGCAACATCCCGAGGCTGCGCAGAGACATGATGTGT 180
|||||
291 TTGATATGATGATATGCGCAACATCCCGAGGCTGCGCAGAGACATGATGTGT 350
181 AACCAAAATGGGGGATTTATGATTCCTCCGCAACACCTGTGTATCGAGGCGCTAC 240
|||||
351 AACCAAAATGGGGGATTTATGATTCCTCCGCAACACCTGTGTATCGAGGCGCTAC 410
241 TCGAACCCCTACTGACACCCCTACTGACAGTCCGATACCCAGAGAGTCCCACTCTCA 300
|||||
411 TCGAACCCCTACTGACACCCCTACTGACAGTCCGATACCCAGAGAGTCCCACTCTCA 470
301 GCTCCAAATATCCACGATCTCCAGGCGCTTATATGCGCCTTGTGATACCATGAT 360
|||||
471 GCTCCAAATATCCACGATCTCCAGGCGCTTATATGCGCCTTGTGATACCATGAT 530
361 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAACAGATTCCTCCAGTGCACACCC 420
|||||
531 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAACAGATTCCTCCAGTGCACACCC 590
421 ACCGAGATCTGATCAATGATGAGAGGCGGTACACCTGCTCTGACACGAGGATTTGG 480
|||||
591 ACCGAGATCTGATCAATGATGAGAGGCGGTACACCTGCTCTGACACGAGGATTTGG 650
481 CTTCTGGGAAGGCGGCTTACATGATGATGATGATGATGATGATGATGATGATGATGAT 540
|||||
651 CTTCTGGGAAGGCGGCTTACATGATGATGATGATGATGATGATGATGATGATGATGAT 710
541 TGTGGGAATGTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 599
|||||
711 TGTGGGAATGTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 770
600 GGATGGAAGGCTTGGCAAGATGTAAGAGTGT--TGCACACGAGAACCCCTG 650
|||||
771 GGATGGAAGGCTTGGCAAGATGTAAGAGTGTGTGCGCACGAGAACCCCTG 823

RESULT 13

BI251219

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

602994427F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5150191 5',
mRNA sequence.
BI251219.1 GI:14800410
EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 717)
AUTHORS NIH-MGC
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-ri@mail.nih.gov
Tissue Procurement: Lochar Hennighausen Ph.D., Robin Humphreys
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: ILAM11371 row: e column: 08
High quality sequence stop: 713.

FEATURES

source

1. 717
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:5150191"
/issue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-Sport6; Site-1: SalI;
Site-2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lochar Hennighausen/Robin Humphreys,
NIH"

BASE COUNT 172 a 186 c 182 g 177 t

ORIGIN

Query Match 44.1%; Score 592.8; DB 13; Length 717;

Best Local Similarity 89.8%; Pred. No. 2.2e-153;

Matches 636; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

326 GGCCTTTATATGCGGCTTTGATACCAAGATGATGAAGAAACAACCAATGTGTGATGTG 385
|||||
1 GGCCTTTGCTGCTCGCGCTTGGTATGATGATGATGAAGCAACAGTGTGTGATGTG 60
386 ACAGATGTCACAGATTTCCACAGTGCACACCCACAGATCTGATCAATCTGAAG 445
|||||
61 ACAGATGTCACAGATTTCCACAGTGCACACCCACAGATCTGATCAATCTGAAG 120
446 GCGGTCACCTGCTCTGACACGACGATATGCTCTGGAAGGCGAGTGTGACA 505
|||||
121 GAGTTACACCTGCTCTGACACGACGATGCTCTGGAAGGCGAGTGTGACA 180
506 TTGATGATGTCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 565
|||||
181 TTGATGATGTCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
566 CTTGTACATGACACCTGTGTTTACCTCAATGAGATGAGAGTCTTCCAAAGATGTA 625
|||||
241 CTTGTACATGACACCTGTGTTTACCTCAATGAGATGAGAGTCTTCCAAAGATGTA 300
626 ACAGATGTCACAGATTTCCACAGTGCACACCCACAGATCTGATCAATCTGAAG 685
|||||
301 ACAGATGTCACAGATTTCCACAGTGCACACCCACAGATCTGATCAATCTGAAG 360
686 TCTGCGGCTGTGACACAGATATGATGATGATGATGATGATGATGATGATGATGATGAT 745
|||||
361 TCTGCGGCTGTGACACAGATATGATGATGATGATGATGATGATGATGATGATGATGAT 420
746 ACAGATGTCACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 805
|||||
421 ACAGATGTCACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
806 ACTTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 865
|||||
481 ACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540

Db 836 AAA-CTGCTCAACACCTAGCATCTTTCACCTGCCGTGTGACC 879

RESULT 11
 B0930927 870 bp mRNA linear EST 20-AUG-2002
 LOCUS AGENCOURT_8955548 NCI CGAP Mam2 Mus musculus cDNA clone
 DEFINITION IMAGE:6439732 5', mRNA sequence.
 ACCESSION B0930927
 VERSION B0930927.1 GI:22345958
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 870)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM13960 row: d column: 05
 High quality sequence stop: 651.
 Location/Qualifiers
 1..870
 /organism="Mus musculus"
 /strain="FVB/N-3"
 /db_xref="taxon:10090"
 /clone_1lb="NCI_CGAP_Mam2"
 /tissue_type="tumor, biopsy sample"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-Sport6; Site:1: Salt;
 Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"

BASE COUNT 200 a 248 c 208 g 212 t 2 others

ORIGIN

Query Match 46.3% Score 622.6 DB 14; Length 870;
 Best Local Similarity 87.7%; Pred. No. 1.3e-161;
 Matches 713; Conservative 0; Mismatches 96; Indels 4; Gaps 3;

QY 1 ATGCCAGGAATAAAAGGATACCTGTTACCATTTGGCTCTGTTCCAAACCCCT 60
 Db 48 ATGCCAGGATTAATAAGATACCTGTTACCATTTGGCTCTGTTCCAAACCCCT 107

QY 61 GGGAAATGCACAGGACAGTGCACGAATGGCTTTGACCTGGATGCCAGTACAGAGTGT 120
 Db 108 GGGAAATGCACAGGACAGTGCACGAATGGCTTTGACCTGGATGCCAGTACAGAGTGT 167

QY 121 TTAGATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 Db 168 CTGATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 227

QY 181 AACCAAAATGGGGGATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 Db 228 AACCAAAATGGGGGATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 287

QY 241 TCGAACCCCTACACGACCCCTACTAGAGTCCGTAACCAAGAGTCCCAACCACTCTCA 300
 Db 288 TCAAAATCCCTACTCTACTACTAGAGGCGCCATACCAAGAGTCCCAACCACTCTCA 347

QY 301 GCTCAAACTATATCCACGATCTCCAGGCTCTTATATGATGATGATGATGATGATGAT 360

Db 348 GCTTCCAACTAACCACGATTTCAAGGCTCTTGTCTGCGGCTTTGGGTAACAGATGAT 407

QY 361 GAAAGCAACCAATGTGTGATGTGTGACAGATGTGTGCAACAGATTTCCACCAAGTCAACCCC 420
 Db 408 GAAAGCAACCAATGTGTGATGTGTGACAGATGTGTGCAACAGATTTCCACCAAGTCAACCCC 467

QY 421 ACCCAGATCTGATCAATCACTGAAGGCGGTGTACACCTGCTCTGACAGGAGTATTTGG 480
 Db 468 ACCCAGATCTGATCAATCACTGAAGGCGGTGTACACCTGCTCTGACAGGAGTATTTGG 527

QY 481 CTTTGGAAAGCCAGTGTCTTACATGTGATGATGATGATGATGATGATGATGATGATGAT 540
 Db 528 CTTTGGAAAGCCAGTGTCTTACATGTGATGATGATGATGATGATGATGATGATGATGAT 587

QY 541 TGTGGAATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 Db 588 TGTGGAATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 647

QY 601 GATGGAAGGCTTGGCAAGATGTGAACGAGTGTGCCACGAGAACCCCTGCTGCAAAACC 660
 Db 648 GATGGAAGGCTTGGCAAGATGTGAACGAGTGTGCCACGAGAACCCCTGCTGCAAAACC 707

QY 661 TGGCGCAACCTACAGGCTCTTTCATCTGCGCTGTGACCCAGAGATATG- AACTTGAGA 719
 Db 708 TGTGTCAACCTATAGGCTCTTTCATCTGCGCTGTGACCCAGAGATATG- AACTTGAGA 767

QY 720 AGATGCGCTTCA-TTGCAATGATATGAGAGAGTGCAGC--TTCTCTGAGTCTCTGCCA 776
 Db 768 AGATGCGCTTCA-TTGCAATGATATGAGAGAGTGCAGC--TTCTCTGAGTCTCTGCCA 827

QY 777 ACATGATGTGTGAACCAAGCCCGGACATATCTT 809
 Db 828 ACATGATGTGTGAACCAAGCCCGGACATATCTT 860

RESULT 12
 B0718885 835 bp mRNA linear EST 16-JUL-2002
 LOCUS AGENCOURT_8241132 lupskl_sympathetic_trunk Homo sapiens cDNA clone
 DEFINITION IMAGE:6187063 5', mRNA sequence.
 ACCESSION B0718885
 VERSION B0718885.1 GI:21857782
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 835)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM13580 row: h column: 08
 High quality sequence stop: 668.
 Location/Qualifiers
 1..835
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1lb="lupskl_sympathetic_trunk"
 /sex="male"
 /tissue_type="sympathetic trunk"
 /dev_stage="adult, 16 yr"
 /lab_host="DH10B"
 /note="Vector: pCMV-Sport6 (Life Technologies); Site:1:

VERSION	B1453348.1	GI:15244004
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus.	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 796)	
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/.	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)	
JOURNAL	Contact: Robert Strausberg, Ph.D. Email: cga@bs-remail.nih.gov Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys CDNA Library Preparation: Life Technologies, Inc. CDNA library arrayed by: The I.M.A.G.E. Consortium (ILNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov Plate: LHAM11637 row: f column: 18 High quality sequence stop: 796.	
FEATURES	Location/Qualifiers	
SOURCE	1..796	
	/organism="Mus musculus"	
	/strain="CS7BL/6J"	
	/db_xref="taxon:10090"	
	/clone IMAGE:5252369"	
	/clone_11b="NCI-CGAP_Mam5"	
	/tissue_type="tumor, gross tissue"	
	/dev_stage="7 months"	
	/lab_host="DH10B"	
	/note="Organ: mammary; Vector: pCMV-SPOrt6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Robin Humphreys, NIH"	
BASE COUNT	195 a 208 c 197 g 196 t	
ORIGIN		
Query Match	47.6%; Score 639.2; DB 13; Length 796;	
Best Local Similarity	89.2%; Pred. No. 3.1e-166;	
Matches 712; Conservative	0; Mismatches 83; Indels 3; Gaps 2	
Y	354 GATGATTAAGAACCAACCAATGTGTGTGATGTGGACAGATGTGACACAGATTCCACACAGTG 413	
D	1 GATGATTAAGAACCAACCAAGTGTGTGATGTGGACAGATGTGACACAGATTCCACACAGTG 60	
Y	414 CAACCCCCACCAGATCTGCATCAATACTGAAGCGGGTACACCTGTCCTGCACGAGCG 473	
D	61 CAACCCCTACCCAGATCTGTATTCACACTGMAAGAGTTTACACTCTCTGCACCGATGG 120	
Y	474 ATATTGGCTTGTGAAGGCCAGTGTGACATTGATGAAATGTGCTATGTATTAGTCCA 533	
D	121 GTACTGGCTTGTGAAGGCGAGTGGCTTAGATATGATGATGATGATGCTATGTATTAGTCCA 180	
Y	534 GCAGCTGTGTGCGAATGTCTCTGATCTTATTTCTTGTACATGCAACCCGTGGTTTAACT 593	
D	181 GCAGCTGTGTGCAATGTCTTCCAGAGTCTTATCTGTACATGCAACCCGTGGTTTAACT 240	
Y	594 CAAATGAGATGGAAGGCTTGTGCAAGATGTGAAGCATGTGGCAGCGAGAACCCTGGCT 653	
D	241 CAAAGCAGATGGAAGGCTTGTGCAAGATGTGAAGCATGTGGCAGCAACTGGAATACTCCGTGT 300	
Y	654 CAAACCTGCTCAACACACTACGGCTTTTCACTTCGCCGTGTGACCCAGATATGAAT 713	
D	301 TCAGACCTGTGTCAACACACTATGGCTTTTATCTTCGCCGTGTGACCCAGATATGAAT 360	
Y	714 TGAGGAAGATGGCTTCAATTCAGTGTGATGAGCAGATGTGCAGCTTCTCTGAGTTCCTCTG 773	
D	361 TGAGGAAGATGGCTTCACTACCTCAGTGTATGAGCAGATGTGCAGCTTCTCTGAGTTCCTCTG 420	
Y	774 CCACATGAGTGTGTGAACCAAGCCGCGACATATCTTGTGCTCTGCTCCAGGCTACT 833	

Query Match	47.3%	Score 636	DB 13	Length 644
BASE COUNT	146 a	195 c	154 g	149 t
ORIGIN				
FEATURES	<p>source</p> <p>1..644</p> <p>/organism="Homo sapiens"</p> <p>/db_xref="taxon:9606"</p> <p>/clone="NHNBC.cn27c11"</p> <p>/clone_id="Normal Human Trabecular Bone Cells"</p> <p>/sex="Female"</p> <p>/tissue_type="Bone"</p> <p>/cell_type="Trabecular Bone Cells"</p> <p>/lab_host="SURE"</p> <p>/note="Organ: Hip; Vector: pBluescript; Site_1: EcoRI; Library constructed by Dr. Marian Young and Dr. Pamela Gehron Robey (NIDCR)"</p>			
COUNT	<p>row: c column: 11</p> <p>Seq primer: -21M13 forward primer (ABI).</p> <p>location/Qualifiers</p>			
COMMENT	<p>Unpublished (1997)</p> <p>Contact: Libin Jia</p> <p>National Genetics Branch</p> <p>National Human Genome Research Institute</p> <p>10/100101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA</p> <p>Tel: 301-402-4877</p> <p>Fax: 301-496-7157</p> <p>Email: libin@helix.nih.gov</p> <p>DNA Sequencing and analyses by National Institutes of Health Intramural Sequencing Center (NISC).</p>			
JOURNAL	<p>SGAP: The Skeletal Genome Anatomy Project</p> <p>Beckstrom-Sternberg, S.M., Green, E.D., Powell, J.I., Yang, L.M., Robey, P.G., Hotchkiss, R.N. and Franccomano, C.A.</p>			
AUTHORS	<p>1 (bases 1 to 644)</p> <p>Jia, L.B., Young, M.F., Touchman, J.W., Bouffard, G.G., Beckstrom-Sternberg, S.M., Green, E.D., Powell, J.I., Yang, L.M., Robey, P.G., Hotchkiss, R.N. and Franccomano, C.A.</p>			
REFERENCE	<p>1 (bases 1 to 644)</p> <p>Jia, L.B., Young, M.F., Touchman, J.W., Bouffard, G.G., Beckstrom-Sternberg, S.M., Green, E.D., Powell, J.I., Yang, L.M., Robey, P.G., Hotchkiss, R.N. and Franccomano, C.A.</p>			
ORGANISM	<p>human.</p> <p>Homo sapiens</p>			
VERSION	<p>644 bp mRNA linear EST 20-JUN-2002</p>			
KEYWORDS	<p>clone NHNBC.cn27c11 random, mRNA sequence.</p>			
LOCUS	<p>BG938621</p>			
DEFINITION	<p>cn27c11.x1 Normal Human Trabecular Bone Cells Homo sapiens cDNA</p>			
RESULT 9	<p>BG938621</p>			
Db	<p>1133 ATTCATTTCGAGATCA 1150</p>			
Oy	<p>779 ATTACATTTCGAGATCA 796</p>			
Db	<p>1073 GAGCCTCCGTTCCCGCGTACATCTTCCAANTGCAAGCACAGACCCTACCCGGGCGCT 1132</p>			
Oy	<p>721 GACCTCCGTTCCCGCGTACATCTTCCAANTGCAAGCAA--AACCCGATACCTGTGCTCT 778</p>			
Db	<p>1013 ACCCTGCGTCGAGACCAAGCCCTTTACCATCTTTGTACCGGGACATGAGCGTGTGAC 1072</p>			
Oy	<p>661 ACACCACTGCGAGAGACCAAGCATTCACATCCCTGTATCGGGACATGAGTGTGTGAC 720</p>			
Db	<p>954 CCGCTGTAGAGAGCCTTATCTGAGAGATCAGTATACCGCTGTATGTCTCTGTAG-A 1012</p>			
Oy	<p>601 CAGCTGTAGAGAGCCTTATCTGCGATGTGTGGAACACCGCTGTATGTCTCTGTAGCA 660</p>			
Db	<p>894 GTGCAACCTGAGAGAGCGTCTACAAATTTCAGAGGGGGCTTCAAATGCAATGACCCCAT 953</p>			
Oy	<p>541 GTGTAACCTGATGAGAGACTGTGTACAACTTACAAAGGGGGCTTCAAATGTAATGATCCAT 600</p>			
Db	<p>481 CCTGTGGATGATTAACCGAAGCTCCGAGGATATCAATGAATGTGAGCACCGAACAAC 540</p>			
Oy	<p>834 CTGCTGGATGACACCGAAGCTCCCAAGACATCAAGATGTGAGCAGACAGAAACAAC 893</p>			
Db	<p>421 TCAACAGAGTGTGAACCAAGCCGGGCTCAATTCTGCTGCGTCCCTCCAGGCTACGT 480</p>			

```

Db      185 GAGTCGACGCTTCTGTGAGTCTCTGTGCAACATGATGTGTGAACACCGCCGACATATAC 244
QY      808 TTCTGCTCCCTGCGCCCTCAGAGCTACATCTCTGTGATGATGACCAACGAGCTGCCAAGACATC 867
Db      245 TTCTGCTCCCTGCGCCCTCAGAGCTACATCTCTGTGATGATGACCAACGAGCTGCCAAGACATC 304
QY      868 AACGAATGTGAGCAGACAGGAACCAACAGCTGCAACCTGACAGACAGCTGTCTACAAATTTACAA 927
Db      305 AACGAATGTGAGCAGACAGGAACCAACAGCTGCAACCTGACAGACAGCTGTCTACAAATTTACAA 364
QY      928 GGGGGGTTAAATGATGATGACACCCCATCCGCTGTGAGAGAGCTTATCTGAGATCATGTAT 987
Db      365 GGGGGGTTAAATGATGATGACACCCCATCCGCTGTGAGAGAGCTTATCTGAGATCATGTAT 424
QY      988 AACCCGCTATGATGCTCTGTGAGAAACCTGTGCTGACAGACACGACCTTTACATCTTG 1047
Db      425 AACCCGCTATGATGCTCTGTGAGAAACCTGTGCTGACAGACACGACCTTTACATCTTG 484
QY      1048 TACCGGGACATGAGAGCTGTGTGACAGAGCTCCGTTCCCGCTGACATCTTCCAAATGCA 1107
Db      485 TACCGGGACATGAGAGCTGTGTGACAGAGCTCCGTTCCCGCTGACATCTTCCAAATGCA 544
QY      1108 GCCAGACACCCGCTGAGCTGTGAGAGAGCTCCGTTCCCGCTGACATCTTCCAAATGCA 1165
Db      545 GCCAGACACCCGCTGAGCTGTGAGAGAGCTCCGTTCCCGCTGACATCTTCCAAATGAGG 603
QY      1166 GCAGAGAAATTTTACATGCGGCAACAGGCGCCCATCATGACACCTGCTGTGATGACACGCC 1225
Db      604 GCAGAGAAATTTTACATGCGGCAACAGGCGCCCATCATGACACCTGCTGTGATGACACGCC 663
QY      1226 CCATCAAA-GGGCCCCGGGAAATCCAGCTGTGAGTGTGAAATGATGATCTGCAACACTGTC 1284
Db      664 CCATCAAAAGGGGGCGCGGTAAATCCAGCTGTGAGTGTGAAATGATGATCTGCAACACTGTC 1342
QY      1285 ATCAACTTCAGAG 1297
Db      722 ATCAAAATTCAAAG 734

```

RESULT 7
 LOCUS B1825121 882 bp mRNA linear EST 04-OCT-2001
 DEFINITION 603038662F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5179550 5',
 mRNA sequence.
 ACCESSION B1825121
 VERSION B1825121.1 GI:15936671
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 882)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 plate: L1447 row: 1 column: 15
 High quality sequence start: 14
 High quality sequence stop: 844.
 Location/Qualifiers
 1..882
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5179550"
 /clone_lib="NIH_MGC_115"

```

/lab_host="DH10B"
/Note="Organ: pooled brain, lung, testis; Vector:
pcmv-sport6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH-MGC Library."
BASE COUNT 191 a 277 c 216 g 197 t 1 others
ORIGIN
Query Match 47.7% Score 641.2; DB 13; Length 882;
Best Local Similarity 98.2%; Pred. No. 9.2e-167;
Matches 690; Conservative 0; Mismatches 9; Indels 4; Gaps 4;

```

1 ATGCCAGAAATTAAGATATCTGATTTACATCTGTGCTCTGTCTTCCAAAGCCCT 60
 179 ATGCCAGAAATTAAGATATCTGATTTACATCTGTGCTCTGTCTTCCAAAGCCCT 238
 61 GGGAAATGACAGGACAGTGTGACAGATGGCTTTGACCTGTGATGCCAGACAGATGT 120
 239 GGGAAATGACAGGACAGTGTGACAGATGGCTTTGACCTGTGATGCCAGACAGATGT 298
 121 TTGATATTTGATGATATCCGAAACATCCCGAGGCTCCGAGAGACATGATGTGT 180
 299 TTGATATTTGATGATATCCGAAACATCCCGAGGCTCCGAGAGACATGATGTGT 358
 181 AACCAAAATGGGGGATTTATGATTTCCCGAGACAAACCTGTGTATGAGGGCCCTAC 240
 359 AACCAAAATGGGGGATTTATGATTTCCCGAGACAAACCTGTGTATGAGGGCCCTAC 418
 241 TCGAACCCCTACTGACAGCTTACTGATGCTGATCCGACAGACAGTGTGCTGATCTCA 300
 419 TCGAACCCCTACTGACAGCTTACTGATGCTGATCCGACAGACAGTGTGCTGATCTCA 478
 301 GCTCAAAATATCCAGATCTCCAGGCTCTTATATCCGCTTGTGATACAGATGAT 360
 479 GCTCAAAATATCCAGATCTCCAGGCTCTTATATCCGCTTGTGATACAGATGAT 538
 361 GAAAGCAACCAATGTGTGATGTGACAGATGTGACAGATTTCCACAGTGTCAACCC 420
 539 GAAAGCAACCAATGTGTGATGTGACAGATGTGACAGATTTCCACAGTGTCAACCC 598
 421 ACCCAATGTGATATGATGAGAGGGGGGTACACCTGCTCTGACAGGAGATATGG 480
 599 ACCCAATGTGATATGATGAGAGGGGGGTACACCTGCTCTGACAGGAGATATGG 658
 481 CTCTGGAAGGCGAGCTTAGACATTTGATGATGCTGTATGTTACTGCGCAGAGCTC 540
 659 CTCTGGAAGGCGAGCTTAGACATTTGATGATGCTGTATGTTACTGCGCAGAGCTC 718
 541 TGTGCGAATTTCTGATCTTATTTGATGATGATGATGATGATGATGATGATGATGAT 599
 719 TGTGCGAATTTCTGATCTTATTTGATGATGATGATGATGATGATGATGATGATGAT 778
 600 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 657
 779 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 838
 658 ACCTGGGTCAACACTTACAGCTTTTATGATGATGATGATGATGATGATGATGATGAT 700
 839 AACTGGGTCAACACTTACAGCTTTTATGATGATGATGATGATGATGATGATGATGAT 880

```

RESULT 8
LOCUS B1453348 796 bp mRNA linear EST 21-AUG-2001
DEFINITION 60313067F1 NC1_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5252369 5',
mRNA sequence.
ACCESSION B1453348

```

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: David N. Louis, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLM9508 row: d column: 01
High quality sequence stop: 675.
Location/Qualifiers
1. .891
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4186704"
/clone_lib="NCI CGAP Brn64"
/tissue_type="gliblastoma with EGFR amplification"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: brain; Vector: PCMV-SPORT6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.57 kb. Constructed by Life Technologies. Note: this is a NCI CGAP library."
BASE COUNT 222 a 258 c 230 g 181 t
ORIGIN
Query Match 49.6%; Score 666.8; DB 12; Length 891;
Best Local Similarity 95.1%; Pred. No. 7.3e-174;
Matches 742; Conservative 0; Mismatches 32; Indels 6; Gaps 5;
Db 304 CCNAATATCCAGCATCTCCAGGCGCTTATATGCGCTTGGATTACAGATGATGATA 363
1 CCAAAATATCCAGCATCTCCAGGCGCTTATATGCGCTTGGATTACAGATGATGATA 59
QY 364 AGCAACCAATGTGTGATGTGACAGAGTGTGCAACAGATTCCACAGATGCAACCCACC 423
Db 60 AGCAACCAATGTGTGATGTGACAGAGTGTGCAACAGATTCCACAGATGCAACCCACC 119
QY 424 CAGATCTGATCAATCTGAAGCGGGGTACACCTGCTCTCCACGAGGATATTGGCTT 483
Db 120 CAGATCTGATCAATCTGAAGCGGGGTACACCTGCTCTCCACGAGGATATTGGCTT 178
QY 484 CTGGAAGGCCAGTCTTACATTTGATGATGATGATGATGATGATGATGATGATGATGAT 543
Db 179 CTGGAAGGCCAGTCTTACATTTGATGATGATGATGATGATGATGATGATGATGATGAT 238
QY 544 GCGAATGTTCTGAGATCTTATTTGATGATGATGATGATGATGATGATGATGATGATGAT 603
Db 239 GCGAATGTTCTGAGATCTTATTTGATGATGATGATGATGATGATGATGATGATGATGAT 298
QY 604 GGAAGGCTTGGCAAGATGAGAGAGTGTGCAACGAGAACCCCTGCGTGAACCTGCG 663
Db 299 GGAAGGCTTGGCAAGATGAGAGAGTGTGCAACGAGAACCCCTGCGTGAACCTGCG 358
QY 664 GTCAACACCTACGGCTTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGAGAGAT 723
Db 359 GTCAACACCTACGGCTTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGAGAGAT 418
QY 724 GCGGCTTCAATGAGATATGAGAGAGTGTGCAACGAGTCTTCTGCGCAATGAG 783
Db 419 GCGGCTTCAATGAGATATGAGAGAGTGTGCAACGAGTCTTCTGCGCAATGAG 478
QY 784 TGTGTGAACCAAGCTGCGGACATATCTGTGCTGCGCTGCGGCTGCGGCTGCGGCTGCG 843
Db 479 TGTGTGAACCAAGCTGCGGACATATCTGTGCTGCGCTGCGGCTGCGGCTGCGGCTGCG 538
QY 844 GACAACCAAGCTGCGGACATATCTGTGCTGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 903
Db 539 GACAACCAAGCTGCGGACATATCTGTGCTGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 598
QY 904 CAGGAGAGCTGTGATCAATTTACAGAGGGGCTTCAAA-TGCATCGACCCCATCCGCTGTGA 962

Db 599 CAGCAGAGCTGCTCAATTTACAGAGGGGGCTTCAAGATGATGACCCATCGCTGTGA 658
QY 963 GAGGCTTATCTGAGATCACTGATTAACCGCT-GTATGTCTCTCTGAGAAACCTTGCT 1021
Db 659 GAGGCTTATCTGAGATCACTGATTAACCGCTGAGATGTCTCTCTGAGAAACCTTGCT 716
QY 1022 GCAGAGACCAAGCTTACATCTTGTACCGGAGATGAGAGCTGTGTCAGAGAGCTCCG 1081
Db 717 GCGAACAACAGCCCTTACATCTGTGTCGCCGACATGCGCCCTGCGGAGAGAGCCGCGCG 776
RESULT 6
LOCUS BG109391 822 bp mRNA linear EST 30-JAN-2001
DEFINITION 602280448F1 NIH_MGC_86 Homo sapiens CDNA clone IMAGE:4368227 5',
mRNA sequence.
ACCESSION BG109391.1 GI:12602897
VERSION BG109391.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 822)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLM10021 row: 0 column: 12
High quality sequence stop: 676.
Location/Qualifiers
1. .822
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4368227"
/clone_lib="NIH_MGC_86"
/tissue_type="osteosarcoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: bone; Vector: PCMV-SPORT6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally. Oligo-dt primed. Average insert size 1.53 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH-MGC library."
BASE COUNT 204 a 227 c 210 g 181 t
ORIGIN
Query Match 49.5%; Score 665; DB 12; Length 822;
Best Local Similarity 97.8%; Pred. No. 2.2e-173;
Matches 717; Conservative 0; Mismatches 10; Indels 6; Gaps 4;
Db 568 TGTACATGCAACCTGCTGTTTACCCCTCAATGAGATGAGAGTCTTGCCAAAGTGTAAAC 627
Db 5 TGTACATGCAACCTGCTGTTTACCCCTCAATGAGATGAGAGTCTTGCCAAAGTGTAAAC 64
QY 628 GAGTGTGCAACCGGAGAACCCCTGCGTGAACCTGCGTCAACACTTACGCTCTTTCATC 687
Db 65 GAGTGTGCAACCGGAGAACCCCTGCGTGAACCTGCGTCAACACTTACGCTCTTTCATC 124
QY 688 TGCCGCTGAGACCCGAGATGATGACTTGTGGAAGATGGGCTTCATGAGTATGAGAC 747
Db 125 TGCCGCTGAGACCCGAGATGATGACTTGTGGAAGATGGGCTTCATGAGTATGAGAC 184
QY 748 GAGTGTGCAACCTTCTGAGTCTCTGCAACATGAGTGTGTGACAGCCGCGACATAC 807

OY		568	TGTCACATGCACAACCTGGTTTAACTCCATAAGTAGGATGGAAAGCTCTGCCAAGATGTGAC	627
Dd		481	TGTACATGCACAACCTGGTTTAACTCCATAAGTAGGATGGAAAGCTCTGCCAAGATGTGAC	540
OY		628	GAGTGTCGCACCGAGAACCCTGGCTCAAACTCGCTCAACACACTACGGCTCTTTGATC	687
Dd		541	GAGTGTCGCACCGAGAACCCTGGCTCAAACTCGCTCAACACACTACGGCTCTTTGATC	600
OY		688	TGCCGCTGTGACCCAGATATGAACCTTGAGAGACATGGCGTTTCATTGCACTGATATGAC	747
Dd		601	TGCCGCTGTGACCCAGATATGAACCTTGAGAGACATGGCGTTTCATTGCACTGATATGAC	660
OY		748	GAGTGCAAGCTTCTGTGAGTTCTCTGCGCAACATGATGTTGTAACCAAGCCGGGACATAC	807
Dd		661	GAGTGCAAGCTTCTGTGAGTTCTCTGCGCAACATGATGTTGTAACCAAGCCGGGACATAC	720
OY		808	TTTGTGCTCTGCGCTCCACAGCTACAT-CTCTGTGATGATCAACG-GAAGCTGCCAAGA--	863
Dd		721	TTTTGCTCTGCGCTCCACAGCTACATCTCTGTGATGATCAACGGAAGCTGCCAAGAAC	780
OY		864	CATCAACGAATGTGAGCACAGAGAACCAACAGCTGCA--CCTACAGACAGCGGTACATTT	922
Dd		781	ATCAACCGGAATGTGAGCACAGAGAACCAACAGCTGCAACCTCTACACAGCGGTACATTT	840
OY		923	T---ACAAGGGGCTTCAATCATGATGACCCCATCGCTGTGAGAG	966
Dd		841	TTACACAGGGGGGCTTCAATCATGATGACCCCATCGCTGTGAGAG	887
RESULT 4				
BMS545444				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
JOURNAL				
COMMENT				

568 TGTACATGCACAACCTGGTTTAACTCCATAAGTAGGATGGAAAGCTCTGCCAAGATGTGAC 627
 Dd 481 TGTACATGCACAACCTGGTTTAACTCCATAAGTAGGATGGAAAGCTCTGCCAAGATGTGAC 540
 OY 628 GAGTGTCGCACCGAGAACCCTGGCTCAAACTCGCTCAACACACTACGGCTCTTTGATC 687
 Dd 541 GAGTGTCGCACCGAGAACCCTGGCTCAAACTCGCTCAACACACTACGGCTCTTTGATC 600
 OY 688 TGCCGCTGTGACCCAGATATGAACCTTGAGAGACATGGCGTTTCATTGCACTGATATGAC 747
 Dd 601 TGCCGCTGTGACCCAGATATGAACCTTGAGAGACATGGCGTTTCATTGCACTGATATGAC 660
 OY 748 GAGTGCAAGCTTCTGTGAGTTCTCTGCGCAACATGATGTTGTAACCAAGCCGGGACATAC 807
 Dd 661 GAGTGCAAGCTTCTGTGAGTTCTCTGCGCAACATGATGTTGTAACCAAGCCGGGACATAC 720
 OY 808 TTTGTGCTCTGCGCTCCACAGCTACAT-CTCTGTGATGATCAACG-GAAGCTGCCAAGA-- 863
 Dd 721 TTTGTGCTCTGCGCTCCACAGCTACATCTCTGTGATGATCAACGGAAGCTGCCAAGAAC 780
 OY 864 CATCAACGAATGTGAGCACAGAGAACCAACAGCTGCA--CCTACAGACAGCGGTACATTT 922
 Dd 781 ATCAACCGGAATGTGAGCACAGAGAACCAACAGCTGCAACCTCTACACAGCGGTACATTT 840
 OY 923 T---ACAAGGGGCTTCAATCATGATGACCCCATCGCTGTGAGAG 966
 Dd 841 TTACACAGGGGGGCTTCAATCATGATGACCCCATCGCTGTGAGAG 887

BMS545444 954 bp mRNA linear EST 20-FEB-2002
 AGENCOURT 6500472 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5729596
 5', mRNA sequence.
 BMS545444
 BMS545444.1 GI:18777513
 EST.

human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 954)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs.remall.nih.gov
 Tissue Procurement: Invitrogen
 cDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMLD)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LMLD at:
 http://image.llnl.gov
 Plate: LLM12727 row: C column: 05
 High quality sequence stop: 676.

Location/Qualifiers
 1..954
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5729596"
 /clone_1lb="NIH_MGC_124"
 /tissue_type="hippocampus"
 /lab_host="DH10B"
 /note="Organ: Brain; Vector: PCMV-SPORT6; Site_1: EcoRV
 (destroyed); Site_2: NotI; RNA source male hippocampus,
 age 27. Library is oligo-dT primed and directionally
 cloned (EcoRV site is destroyed upon cloning). Average
 insert size 1.4 kb, insert size range 0.9-4 kb. Library is
 normalized and enriched for full-length clones and was
 constructed by C. Gruber (Invitrogen). Research Genetics
 tracking code 012."

PAGE COUNT 232 a 278 c 238 g 206 t

ORIGIN	Query Match	57.48; Score 771; DB 13; Length 954;
	Best Local Similarity 100.0%; Pred. No. 9.2e-203;	
	Matches 771; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	574 TGCACACCTGGTTTATACCTCAATGAGGATGGAAAGGCTTGGCCAAAGATGTGAACAGAGT	633
Db	8 TGCACACCTGGTTTATACCTCAATGAGGATGGAAAGGCTTGGCCAAAGATGTGAACAGAGT	67
OY	634 GCCACCGAGAACCCCTGGCTGTCGAACACCTGCTCAACACACTACGAGCTCTTTTCATCTGC	693
Db	68 GCCACCGAGAACCCCTGGCTGTCGAACACCTGCTCAACACACTACGAGCTCTTTTCATCTGC	127
OY	694 TGTGACCCAGATATGAACTTGAGGAGATGGCTTCACTTTCAGTATATGACAGATG	753
Db	128 TGTGACCCAGATATGAACTTGAGGAGATGGCTTCACTTTCAGTATATGACAGATG	187
OY	754 AGCTTCTCTGAGTTCCCTCTGCCCAACATGATGTGTGAACACGACCGGACATACCTTCTGC	813
Db	188 AGCTTCTCTGAGTTCCCTCTGCCCAACATGATGTGTGAACACGACCGGACATACCTTCTGC	247
OY	814 TCTTGCCTCCAGGCTACATCTCTGCTGATGACAAACGGAAGCTGCGAAGATCAAGAA	873
Db	248 TCTTGCCTCCAGGCTACATCTCTGCTGATGACAAACGGAAGCTGCGAAGATCAAGAA	307
OY	874 TGTGAGCAGAGAACACACCTGGAACCTTGAGACAGACGCTCAATTTCAAGAGGGGC	933
Db	308 TGTGAGCAGAGAACACACCTGGAACCTTGAGACAGACGCTCAATTTCAAGAGGGGC	367
OY	934 TTCAAATGATGACACCCATCCGCTGTGAGGAGCCCTTATCTGAGATCATGATAACCG	993
Db	368 TTCAAATGATGACACCCATCCGCTGTGAGGAGCCCTTATCTGAGATCATGATAACCG	427
OY	994 TGTATGTGCTCGTGAGAACCTCGGCTGCGAGAGACGACCCCTTACCATCTTGAACCG	1053
Db	428 TGTATGTGCTCGTGAGAACCTCGGCTGCGAGAGACGACCCCTTACCATCTTGAACCG	487
OY	1054 GACATGAGACGTGTGTCTAGAGAGCGCTCCGCTCCGCTGACATCTTCCAATGCAAGCCACG	1113
Db	488 GACATGAGACGTGTGTCTAGAGAGCGCTCCGCTCCGCTGACATCTTCCAATGCAAGCCACG	547
OY	1114 ACCCGCTACCTGGGGCTATTACATTTTCCAGATCAAAATTTGGGAATGAGGGCAGAGAA	1173
Db	548 ACCCGCTACCTGGGGCTATTACATTTTCCAGATCAAAATTTGGGAATGAGGGCAGAGAA	607
OY	1174 TTTTACATGCGGCAAGAGGGGCCCATCATGCTCCACCCGATGATGACAGGCCCATCAAA	1233
Db	608 TTTTACATGCGGCAAGAGGGGCCCATCATGCTCCACCCGATGATGACAGGCCCATCAAA	667
OY	1234 GGGCCCCGGGAATTCACAGCTGGACTTGGAAATGATCACTGTTCACACTGTCAATCTTC	1293
OY	1294 AGAGGAGAGCTCGTATGATCGAGCTCGGATATATGATGATGCGAGTACCCATTC	1344
Db	668 GGGCCCCGGGAATTCACAGCTGGACTTGGAAATGATCACTGTTCACACTGTCAATCTTC	727
Db	728 AGAGGAGAGCTCGTATGATCGAGCTCGGATATATGATGATGCGAGTACCCATTC	778
RESULT 5	BF339448	891 bp mRNA linear EST 22-NOV-2000
LOCUS	BF339448	60203910221 NCL_CGAP_Brn64 Homo sapiens cDNA clone IMAGE:4186704
DEFINITION	5', mRNA sequence.	
VERSION	BF339448	
KEYWORDS	BF339448.1 GI:11285903	
SOURCE	EST.	
ORGANISM	human.	
	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
REFERENCE	1 (bases 1 to 891)	
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/ .	

[illegible]

FEATURES	source
<p>NIH-MGC http://mgc.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabbs-r@mail.nih.gov Tissue Procurement: Dr. James R. Lupski cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLM13581 row: c column: 21 High quality sequence stomp: 646.</p>	<p>Location/Qualifiers 1. 904 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:6187340" /clone_lib="lupski_sympathetic_trunk" /sex="male" /lisse_type="sympathetic trunk" /dev_stage="adult, 16 yr" /lab_host="DH10B" /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1: NotI; Site_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TGGACCCAGGGGTCG-3' and 5'-GACTAGTCTAGATCGAGCGGCCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.9 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine); available through Life Technologies."</p>
BASE COUNT	219 a 264 c 219 g 202 t
ORIGIN	
Query Match	58.9%; Score 792.2; DB 14; Length 904;
Best Local Similarity	97.1%; Pred. No. 1.2e-208;
Matches	861; Conservative 0; Mismatches 18; Indels 8; Gaps 5;
QY	88 GGCCTTGACCTGATCGCCAGTCAGACAGTGTATGATTTGATGATGCGGAACCATC 147
Db	1 GGCCTTGACCTGATCGCCAGTCAGACAGTGTATGATTTGATGATGCGGAACCATC 60
QY	148 CCGGAGGCGTCCGAGAGAGACATATGTGTATTAACCAAAATGGCGGATTTATGCAAT 207
Db	61 CCGGAGGCGTCCGAGAGAGACATATGTGTATTAACCAAAATGGCGGATTTATGCAAT 120
QY	208 CCCCAGCAAAACCTGTGTATCGAGGGCCCTACTACGAAACCCCTACTGACCCCTACTCA 267
Db	121 CCCCAGCAAAACCTGTGTATCGAGGGCCCTACTACGAAACCCCTACTGACCCCTACTCA 180
QY	268 GGTCCGTAACCAAGCAGCTGCCCAACACTCTCAGCTCCAAACTATGCCAGATCTCAGG 327
Db	181 GGTCCGTAACCAAGCAGCTGCCCAACACTCTCAGCTCCAAACTATGCCAGATCTCAGG 240
QY	328 CCTTTATATGCGCCTTTGATATCCATATGATGATTAAGAAACCAATGTGTGATGTGGAC 387
Db	241 CCTTTATATGCGCCTTTGATATCCATATGATGATTAAGAAACCAATGTGTGATGTGGAC 300
QY	388 GAGTGTCAACAGATTTCCACCAAGTCGAACCCCAACCCAGATCTGCATCAATACGAAGGC 447
Db	301 GAGTGTCAACAGATTTCCACCAAGTCGAACCCCAACCCAGATCTGCATCAATACGAAGGC 360
QY	448 GGGTACACCTGTCTCTGCACCAAGGATATTTGGCTTCTGGAAGGCCAGTCTTAGACATT 507
Db	361 GGGTACACCTGTCTCTGCACCAAGGATATTTGGCTTCTGGAAGGCCAGTCTTAGACATT 420
QY	508 GATGAATGTGCGATATGTTACTGTCCGACGACCTGTGCGCAATGTCTCGATCTTATCT 567
Db	421 GATGAATGTGCGATATGTTACTGTCCGACGACCTGTGCGCAATGTCTCGATCTTATCT 480

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5752785"
/clone_1ib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
PCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH MGC Library."

BASE COUNT      235 a      300 c      246 g      233 t
ORIGIN
Query Match      64.4%; Score 866; DB 14; Length 1014;
Best Local Similarity 98.4%; Pred. No. 4.2e-229;
Matches 917; Conservative 0; Mismatches 10; Indels 5; Gaps 4;

1 ATGCCAGAAATAAAGATACACTGTTACCATTTGCTCTCTGCTTCCAAAGCCCT 60
|||||
88 ATGCCAGAAATAAAGATACACTGTTACCATTTGCTCTCTGCTTCCAAAGCCCT 147
|||||
61 GGGAAATGCACAGGACAGTGCAGCAATGGCTTGACCTGGATGCCAGTGCAGACAGTGT 120
|||||
148 GGGAAATGCACAGGACAGTGCAGCAATGGCTTGACCTGGATGCCAGTGCAGACAGTGT 207
|||||
121 TTAGATATTGATGATATGCCGAACCATCCCGAGGCTGCGGAGAGACATGATGTGT 180
|||||
208 TTAGATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 267
|||||
181 AACCAAAATGGGGGATTTATGATCCCGGACAAACCTGTTATCGAGGGGCGCCAC 240
|||||
268 AACCAAAATGGGGGATTTATGATCCCGGACAAACCTGTTATCGAGGGGCGCCAC 327
|||||
241 TCGAACCCTTACTGACCCCTTACTGACGTCGCTGACCTGACGTCGCTGACCTGAC 300
|||||
338 TCGAACCCTTACTGACCCCTTACTGACGTCGCTGACCTGACGTCGCTGACCTGAC 387
|||||
301 GCTCCAAATATCCACGATCTCCAGGCTCTTATATGCGGCTTGGATACCAATGAT 360
|||||
388 GCTCCAAATATCCACGATCTCCAGGCTCTTATATGCGGCTTGGATACCAATGAT 447
|||||
361 GAAAGCAACCAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
|||||
448 GAAAGCAACCAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 507
|||||
421 ACCCAGATCTGATCAATATGTAAGCGGGGTACACCTGCTCTGACCGAGGATATTGG 480
|||||
508 ACCCAGATCTGATCAATATGTAAGCGGGGTACACCTGCTCTGACCGAGGATATTGG 567
|||||
481 CTTTGGGAAGGCGCAGTGTAGACATTTGATGATGATGATGATGATGATGATGATGAT 540
|||||
568 CTTTGGGAAGGCGCAGTGTAGACATTTGATGATGATGATGATGATGATGATGATGAT 627
|||||
541 TGTGGGAATGTCCTGGATCTCTTCTTGTATGATGATGATGATGATGATGATGATGAT 600
|||||
628 TGTGGGAATGTCCTGGATCTCTTCTTGTATGATGATGATGATGATGATGATGATGAT 687
|||||
601 GATGGAAGGCTTGTCCAGATGTGAAGCATGTGCCACGAGAACCCCTGGTGCACAAAC 660
|||||
688 GATGGAAGGCTTGTCCAGATGTGAAGCATGTGCCACGAGAACCCCTGGTGCACAAAC 747
|||||
661 TGGCTCAACACCTAGCGCTCTTCACTGCGCGCTGTGACCCAGAGATATGAATTGAGGAA 720
|||||
748 TGGCTCAACACCTAGCGCTCTTCACTGCGCGCTGTGACCCAGAGATATGAATTGAGGAA 807
|||||
721 GATGCGCTTCACTGCGCGCTGTGACCCAGAGATATGAATTGAGGAA 780
|||||
808 GATGCGCTTCACTGCGCGCTGTGACCCAGAGATATGAATTGAGGAA 867

```

```

OY 781 GACTGTGTGAACACGCGCGGACACATCTTCTGCTCTGCTCCAGGCTACATCTGCTG 840
|||||
DB 868 GACTGTGTGAACACGCGCGGACACATCTTCTGCTCTGCTCCCT-CAGGCTACATCTGCTG 926
|||||
OY 841 GATGACAAACGGAAGCTGCCAAGACATCAAGCAATGTGAGCAGACAGAACACGTCGAC 900
|||||
DB 927 AATGACAAACGGAAGCTGCCAAGACATC-ACGAATGTGAGCAGACAGAACACACAC--TGCAC 983
|||||
OY 901 CTGACAGACAGCTGCTTCAATTTTACAGGGGG 932
|||||
DB 984 CTGACAGACAG-GTGCTCAATTTTACAGGGGG 1014
|||||

RESULT 2
B0720516      996 bp      mRNA      linear      EST 16-JUL-2002
LOCUS
DEFINITION
AGENCOURT_8219503 lupski_sym pathetic_trunk Homo sapiens cDNA clone
IMAGE:6188164 5', mRNA sequence.
B0720516
B0720516.1 GI:21859413
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 996)
NIH-MGC http://mgi.ncl.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM3583 row: f column: 05
High quality sequence stop: 634.
Location/Qualifiers
1. 996
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6188164"
/clone_1ib="lupski_sym pathetic_trunk"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TGACAGTCTTGTAGATGCGAGGCGCGCTT(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."

BASE COUNT      245 a      284 c      242 g      224 t      1 others
ORIGIN
Query Match      61.4%; Score 825.4; DB 14; Length 996;
Best Local Similarity 97.9%; Pred. No. 7.6e-218;
Matches 889; Conservative 0; Mismatches 12; Indels 7; Gaps 5;

OY 51 TCCAAACCCCGGGAATGCACAGGACAGTGCAGCAATGGCTTTGACCTGATCGCCACTC 110
|||||
DB 1 TTCCAGCCCTCGGGAATGCACAGGACAGTGCAGCAATGGCTTTGACCTGATCGCCACTC 60
|||||
OY 111 AGACAGTGTTTAGATATTGATGATGATGATGATGATGATGATGATGATGATGATGAT 170
|||||

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 10:48:06 ; Search time 1982.41 Seconds
(without alignments)
10979.953 Million cell updates/sec

Title: US-09-674-379a-11

Perfect-score: 1344
Sequence: 1 atgcagcagataaaagat.....atgttcagcagatccattc 1344

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estlu:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	866	64.4	1014	14	BM921371 AGENCOURT
2	825.4	61.4	996	14	BQ720516 AGENCOURT
3	792.2	58.9	904	14	BQ716569 AGENCOURT
4	771	49.4	954	13	BM545444 AGENCOURT
5	666.8	49.6	891	12	BF339448 AGENCOURT
6	665	49.5	822	12	BG109391 BG109391 602280448

7	641.2	47.7	882	13	BI825121	BI825121 603038662
8	639.2	47.6	796	13	BI453348	BI453348 603173067
9	636	47.3	644	13	BG938621	BG938621 cn27c11.x
10	624.6	46.5	880	13	BI764020	BI764020 603043221
11	622.6	46.3	870	14	BQ930927	BQ930927 AGENCOURT
12	619	46.1	835	14	BQ718885	BQ718885 AGENCOURT
13	592.8	44.1	717	13	BI251219	BI251219 602994427
14	582.4	43.3	794	13	BI837271	BI837271 603090563
15	576.4	42.9	933	14	BQ716088	BQ716088 AGENCOURT
16	573.6	42.7	912	13	BF090699	BF090699 QV2-NT104
17	564.4	42.0	627	13	BI768516	BI768516 603057118
18	555	41.3	721	13	BI523975	BI523975 603052129
19	553.2	41.2	815	12	BG174360	BG174360 602336371
20	540.2	40.2	627	13	BI183882	BI183882 UNL-P-FN-
21	537.8	40.0	745	12	BG431026	BG431026 602498636
22	534.4	39.8	1006	9	AL545549	AL545549 AL545549
23	521.2	38.8	762	13	BI834712	BI834712 603090304
24	510	37.9	531	10	AW957458	AW957458 EST369423
25	486	36.2	887	9	AL041285	AL041285 DKFZP434M
26	480.4	35.7	889	14	BQ897144	BQ897144 AGENCOURT
27	479	35.6	491	14	BM707581	BM707581 UI-E-C11-
28	477.8	35.6	494	14	BQ695207	BQ695207 1001350 H
29	456.2	33.9	685	9	AL543659	AL543659 AL543659
30	452	33.6	629	13	BI489863	BI489863 603031610
31	449.2	33.4	1015	13	BM544363	BM544363 AGENCOURT
32	442	32.9	690	12	BF104873	BF104873 601822595
33	434.2	32.3	506	14	BQ723059	BQ723059 AGENCOURT
34	415.4	30.9	576	10	AW142226	AW142226 EST292462
35	409.8	30.5	770	12	BG542457	BG542457 602569534
36	391.2	29.1	680	10	AW106432	AW106432 um27g01.Y
37	389.6	29.0	1101	12	BG245848	BG245848 602358649
38	387.8	28.9	391	14	BM699401	BM699401 UI-E-DX1-
39	384.2	28.6	462	10	BE656312	BE656312 UI-M-BH0-
40	382.4	28.5	719	12	BG718928	BG718928 602692222
41	375.4	27.9	377	14	N95751	N95751 zb55e06.s1
42	371.2	27.6	1850	11	BC018871	BC018871 Homo sapi
43	370.4	27.6	389	10	BE004365	BE004365 CMO-BN010
44	368.8	27.4	858	9	AL546863	AL546863 AL546863
45	366.6	27.3	430	10	AW744572	AW744572 UR32n12.Y

ALIGNMENTS

RESULT 1
LOCUS BM921371 1014 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT_6626255 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5752785
ACCESSION BM921371
VERSION BM921371.1 GI:19371750
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1014)
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rt@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLNL12787 row: 1 column: 10
High quality sequence stop: 738.
Location/Qualifiers
1. 1014

FEATURES

source

```
QY 481 CTTCTGGAAGGCCAGTGTCTTATGACATTTGATGAATGTGCTATGTTACTGACAGACTC 540
   |||||||
Db 930 CTTCTGGAAGGCCAGTGTCTTATGACATTTGATGAATGTGCTATGTTACTGACAGACTC 989
QY 541 TGTGCGAATGTTCCTGATGATCCATTTCTTGTACATGCAACCCGTGTTTACCCCTCAATGAG 600
   |||||||
Db 990 TGTGCGAATGTTCCTGATGATCCATTTCTTGTACATGCAACCCGTGTTTACCCCTCAATGAG 1049
QY 601 GATGGAAGTCTTGTCCCAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCGTGCAACC 660
   |||||||
Db 1050 GATGGAAGTCTTGTCCCAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCGTGCAACC 1109
QY 661 TGGGTCAACACCTAGAGGCTCTTTCATCTGCGCTGTGACCCGAGATATGACTTGAGAA 720
   |||||||
Db 1110 TGGGTCAACACCTAGAGGCTCTTTCATCTGCGCTGTGACCCGAGATATGACTTGAGAA 1169
QY 721 GATGGCTTCAATGACATGATATGAGACAGTGTGACAGCTTCTGTAGTTCCCTGCAACAT 780
   |||||||
Db 1170 GATGGCTTCAATGACATGATATGAGACAGTGTGACAGCTTCTGTAGTTCCCTGCAACAT 1229
QY 781 GAGTGTGAACCAAGCCCGGACATCTTGTCTGCTCTGCTCCCTCCAGGCTACATCCGTCTG 840
   |||||||
Db 1230 GAGTGTGAACCAAGCCCGGACATCTTGTCTGCTCTGCTCCCTCCAGGCTACATCCGTCTG 1289
QY 841 GATGACAACCGAAGCTGCCAAGACATCAACGAATGTGAGACAGGAGACACACAGTGCAC 900
   |||||||
Db 1290 GATGACAACCGAAGCTGCCAAGACATCAACGAATGTGAGACAGGAGACACACAGTGCAC 1349
QY 901 CTGCGACAGACGTGCTACAAATTTACAAGGGGGCTTCAAAATGCATGACCCCATCCGCTGT 960
   |||||||
Db 1350 CTGCGACAGACGTGCTACAAATTTACAAGGGGGCTTCAAAATGCATGACCCCATCCGCTGT 1409
QY 961 GAGGAGCTTATCTGAGAGATCAGATGATACCGCTGATGTGCTGCTGAGAAACCTGGC 1020
   |||||||
Db 1410 GAGGAGCTTATCTGAGAGATCAGATGATACCGCTGATGTGCTGCTGAGAAACCTGGC 1469
QY 1021 TGCAGAGACCGACCTTTACCATCTTGTACCGGGACATGAGAGTGTGCAGAGACGCTCC 1080
   |||||||
Db 1470 TGCAGAGACCGACCTTTACCATCTTGTACCGGGACATGAGAGTGTGCAGAGACGCTCC 1529
QY 1081 GTTCCCGCTGACATCTTCCAAATGCACAGCCGCTACCTGCGGCTATTACATT 1140
   |||||||
Db 1530 GTTCCCGCTGACATCTTCCAAATGCACAGCCGCTACCTGCGGCTATTACATT 1589
QY 1141 TTCCAGATCAAACTGGGAATGAGGGCAGAGAAATTTTACATGCGGCAAAAGGGCCCATC 1200
   |||||||
Db 1590 TTCCAGATCAAACTGGGAATGAGGGCAGAGAAATTTTACATGCGGCAAAAGGGCCCATC 1649
QY 1201 AGTGCCACCTGTGATGACAGGCCCATCAAAAGGGCCCGGGAATCCAGCTGACTTG 1260
   |||||||
Db 1650 AGTGCCACCTGTGATGACAGGCCCATCAAAAGGGCCCGGGAATCCAGCTGACTTG 1709
QY 1261 GAAATGATCACTGTCAACACTGTTCATCAACTTCAGAGGACAGTCCGTGATCCGACTGCGG 1320
   |||||||
Db 1710 GAAATGATCACTGTCAACACTGTTCATCAACTTCAGAGGACAGTCCGTGATCCGACTGCGG 1769
QY 1321 ATATATGTGCGAGTACCCATTG 1344
   |||||||
Db 1770 ATATATGTGCGAGTACCCATTG 1793
```

Search completed: July 3, 2003, 17:53:35
Job time : 221.968 secs

QY 181 AACCAAAATGGGGTATTATGCAATTCCTCCGGACAAACCTGTGTATGAGGGCCCTAC 240
DB 630 AACCAAAATGGGGTATTATGCAATTCCTCCGGACAAACCTGTGTATGAGGGCCCTAC 689
QY 241 TCGAATCCCTACGAGCCCTACGAGTCCGTACCCAGAGCTGCCCCACACTCTCA 300
DB 690 TCGAATCCCTACGAGCCCTACGAGTCCGTACCCAGAGCTGCCCCACACTCTCA 749
QY 301 GCTCCAAATATCCAGATCTCCAGGCTCTTATATGCGGTTTGATACAGATGAT 360
DB 750 GCTCCAAATATCCAGATCTCCAGGCTCTTATATGCGGTTTGATACAGATGAT 809
QY 361 GAAAGCAACCAATGTGTGATGTGAGAGCTGTGCAACAGATTCACCAAGTGCACACCC 420
DB 810 GAAAGCAACCAATGTGTGATGTGAGAGCTGTGCAACAGATTCACCAAGTGCACACCC 869
QY 421 ACCAGATCTGCATCACTAGAGGGGGGTACACCTCTCTCTGACCGAGATATG 480
DB 870 ACCAGATCTGCATCACTAGAGGGGGGTACACCTCTCTCTGACCGAGATATG 929
QY 481 CTCTGGAAGGCGAGCTCTTACATTTGATGATGTGCTATGTTACTGCGAGAGCTC 540
DB 930 CTCTGGAAGGCGAGCTCTTACATTTGATGATGTGCTATGTTACTGCGAGAGCTC 989
QY 541 TGTGCGAATGTTCTGATCTCTATCTTGTATGATGCAACCTGTTTACCTCAATGAG 600
DB 990 TGTGCGAATGTTCTGATCTCTATCTTGTATGATGCAACCTGTTTACCTCAATGAG 1049
QY 601 GATGGAAGTCTTGCCAAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGGGTGCAACC 660
DB 1050 GATGGAAGTCTTGCCAAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGGGTGCAACC 1109
QY 661 TGGGTAAACCTTAACGCTCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGAA 720
DB 1110 TGGGTAAACCTTAACGCTCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGAA 1169
QY 721 GATGGGCTTATGAGTATGAGAGAGTGTGAGAGCTTCTGTGATCTCTGCAACAT 780
DB 1170 GATGGGCTTATGAGTATGAGAGAGTGTGAGAGCTTCTGTGATCTCTGCAACAT 1229
QY 781 GAGTGTGTGAACCGCCGACATATCTTGTCTGCTGCTCCAGGTATCATCTGCTG 840
DB 1230 GAGTGTGTGAACCGCCGACATATCTTGTCTGCTGCTCCAGGTATCATCTGCTG 1289
QY 841 GATGCAACCGAAGCTGTGCAAGACATCAACGAATGTGAGACAGAAACACAGCTGCAAC 900
DB 1290 GATGCAACCGAAGCTGTGCAAGACATCAACGAATGTGAGACAGAAACACAGCTGCAAC 1349
QY 901 CTGCAAGACAGCTGCTACATTTACAAAGGGGGCTTCAAAATGATCGAACCCCATCGCTGT 960
DB 1350 CTGCAAGACAGCTGCTACATTTACAAAGGGGGCTTCAAAATGATCGAACCCCATCGCTGT 1409
QY 961 GAGGAGCTTATCTGAGATCACTGATTAACCGCTGTATGTCTGCTGAGAACCTGTGC 1020
DB 1410 GAGGAGCTTATCTGAGATCACTGATTAACCGCTGTATGTCTGCTGAGAACCTGTGC 1469
QY 1021 TGCAGAGACAGCCCTTACCATCTTGTACCGGGAGATGAGAGTGTGTCAGAGCTGC 1080
DB 1470 TGCAGAGACAGCCCTTACCATCTTGTACCGGGAGATGAGAGTGTGTCAGAGCTGC 1529
QY 1081 GTTCCCGTGAATCTTCCAAATGCAAGCCAGACCGCTTACCTGGGGCTTATTAAT 1140
DB 1530 GTTCCCGTGAATCTTCCAAATGCAAGCCAGACCGCTTACCTGGGGCTTATTAAT 1589
QY 1141 TTCCAGATCAAAATCTGGGAATAGAGGAGAGAAATTTTACATGCGGCAACAGGGCCCATC 1200
DB 1590 TTCCAGATCAAAATCTGGGAATAGAGGAGAGAAATTTTACATGCGGCAACAGGGCCCATC 1649
QY 1201 AGTGCACCTGTGTATGACAGCCCATCAAAAGGGGGGAAATTCAGCTGTGACTTG 1260
DB 1650 AGTGCACCTGTGTATGACAGCCCATCAAAAGGGGGGAAATTCAGCTGTGACTTG 1709
QY 1261 GAAATGATCACTGTCAACACTGTCACTTCAAGAGGAGCTCCGCTGATCCGACTGTGCG 1320

DB 1710 GAAATGATCACTGTCAACACTGTCACTTCAAGAGGAGCTCCGTGTGACTGTGCG 1769
QY 1321 ATATATGTGTGCGAGTACCATTC 1344
DB 1770 ATATATGTGTGCGAGTACCATTC 1793

RESULT 15
US-10-180-552-407
; Sequence 407, Application US/10180552
; Publication No. US20030022300A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zhenlin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C153
; CURRENT APPLICATION NUMBER: US/10/180, 552
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 407
; LENGTH: 2609
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-10-180-552-407

Query Match 99.9%; Score 1342.4; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1343; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCCAGAAATAAAAAGATACCTGTTACATCTGTGCTCTGTCTTCCAGGCCCT 60
DB 450 ATGCCAGAAATAAAAAGATACCTGTTACATCTGTGCTCTGTCTTCCAGGCCCT 509
QY 61 GGGAAATGCACAGGACAGTGCAGAAATGCTTTGACCTGTGATGCGCAGTGCAGAGTGT 120
DB 510 GGGAAATGCACAGGACAGTGCAGAAATGCTTTGACCTGTGATGCGCAGTGCAGAGTGT 569
QY 121 TTAGATATTTGATGAATGCGGAACCATCCCGAGGCTGCGGAGAGACATGATGTGTT 180
DB 570 TTAGATATTTGATGAATGCGGAACCATCCCGAGGCTGCGGAGAGACATGATGTGTT 629
QY 181 AACCAAAATGGGGTATTATGCAATTCCTCCGGACAAACCTGTGTATGAGGGCCCTAC 240
DB 630 AACCAAAATGGGGTATTATGCAATTCCTCCGGACAAACCTGTGTATGAGGGCCCTAC 689
QY 241 TCGAATCCCTACGAGCCCTACGAGTCCGTACCCAGAGCTGCCCCACACTCTCA 300
DB 690 TCGAATCCCTACGAGCCCTACGAGTCCGTACCCAGAGCTGCCCCACACTCTCA 749
QY 301 GCTCCAAATATCCAGATCTCCAGGCTCTTATATGCGGTTTGATACAGATGAT 360
DB 750 GCTCCAAATATCCAGATCTCCAGGCTCTTATATGCGGTTTGATACAGATGAT 809
QY 361 GAAAGCAACCAATGTGTGATGTGAGAGCTGTGCAACAGATTCACCAAGTGCACACCC 420
DB 810 GAAAGCAACCAATGTGTGATGTGAGAGCTGTGCAACAGATTCACCAAGTGCACACCC 869
QY 421 ACCAGATCTGCATCACTAGAGGGGGGTACACCTCTCTCTGACCGAGATATG 480
DB 870 ACCAGATCTGCATCACTAGAGGGGGGTACACCTCTCTCTGACCGAGATATG 929

TYPE: DNA
ORGANISM: Homo Sapien
us-10-176-757-407

Query Match 99.98; Score 1342.4; DB 9; Length 2609;
Best Local Similarity 99.98; Pred. No. 0;
Matches 1343; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
1 ATGCCAGAAATAAAGAGATCTACCTGTTACCATTTCTGCTCTCTCTCCAAAGCCCT 60
450 ATGCCAGAAATAAAGAGATCTACCTGTTACCATTTCTGCTCTCTCTCCAAAGCCCT 509
61 GGAATGACAGACAGACAGTGCAGAAATGCTTACCTGATCCAGTCCAGTGCAGAGTGT 120
510 GGAATGACAGACAGTGCAGAAATGCTTACCTGATCCAGTCCAGTGCAGAGTGT 569
121 TTGATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
570 TTGATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 629
181 AACCAAAATGGGGGATTTATGATGATGATGATGATGATGATGATGATGATGATGATG 240
630 AACCAAAATGGGGGATTTATGATGATGATGATGATGATGATGATGATGATGATGATG 689
241 TCGAACCCCTACTGACAGCCCTACTGACAGTCCGATCCAGCAGAGTCCGATCCAG 300
690 TCGAACCCCTACTGACAGCCCTACTGACAGTCCGATCCAGCAGAGTCCGATCCAG 749
301 GCTCCAACTATCCAGATCTCCAGGCTCTTATATGATGATGATGATGATGATGATGAT 360
750 GCTCCAACTATCCAGATCTCCAGGCTCTTATATGATGATGATGATGATGATGATGAT 809
361 GAAAGAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
810 GAAAGAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 869
421 ACCCAGATCTGATCAATATGATGATGATGATGATGATGATGATGATGATGATGATG 480
870 ACCCAGATCTGATCAATATGATGATGATGATGATGATGATGATGATGATGATGATG 929
481 CTTCTGGAGGCGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
930 CTTCTGGAGGCGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 989
541 TGTGGAATGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
990 TGTGGAATGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1049
601 GATGGAAGTCTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
1050 GATGGAAGTCTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1109
661 TCCGTCACACCTGACGCTCTTCAATGATGATGATGATGATGATGATGATGATGATGAT 720
1110 TCCGTCACACCTGACGCTCTTCAATGATGATGATGATGATGATGATGATGATGATGAT 1169
721 GATGGGCTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
1170 GATGGGCTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1229
781 GAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
1230 GAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1289
841 GATGACAAACCGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
1290 GATGACAAACCGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1349
901 CTGACAGCAGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
1350 CTGACAGCAGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1409
961 GAGGAGCCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
```

```
1410 GAGGAGCCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1469
1021 TGCAGAGACCAAGCCCTTACCATCTTGTACCGGGACATGAGAGTGTGTCAGAGCCTCC 1080
1470 TGCAGAGACCAAGCCCTTACCATCTTGTACCGGGACATGAGAGTGTGTCAGAGCCTCC 1529
1081 GTTCCCGCTGATCATCTTCCAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
1530 GTTCCCGCTGATCATCTTCCAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1589
1141 TTCCAGATCAAAATCTGGGAATGAGGGCAGAGAAATTTTACATCGGCAAAAGGGCCCATC 1200
1590 TTCCAGATCAAAATCTGGGAATGAGGGCAGAGAAATTTTACATCGGCAAAAGGGCCCATC 1649
1201 AGTCCACCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
1650 AGTCCACCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1709
1261 GAAATGATCACTGTCACACAGTGCATCAACTTCAGAGGAGCTCCGATGATGATGATGATG 1320
1710 GAAATGATCACTGTCACACAGTGCATCAACTTCAGAGGAGCTCCGATGATGATGATGATG 1769
1321 ATATATGTGTGCGAGTACCCATTTC 1344
1770 ATATATGTGTGCGAGTACCCATTTC 1793
```

RESULT 14

US-10-176-913-407

Sequence 407, Application US/10176913

Publication No. US20030022298A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Pan, James

APPLICANT: Smith, Victoria

APPLICANT: Watanabe, Colin K.

APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P3430R1C66

CURRENT APPLICATION NUMBER: US/10/176,913

CURRENT FILING DATE: 2002-06-20

Prior Application removed - See file Wrapper or Palm

NUMBER OF SEQ ID NOS: 612

SEQ ID NO 407

LENGTH: 2609

TYPE: DNA

ORGANISM: Homo Sapien

US-10-176-913-407

Query Match 99.98; Score 1342.4; DB 9; Length 2609;

Best Local Similarity 99.98; Pred. No. 0;

Matches 1343; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
1 ATGCCAGAAATAAAGAGATCTACCTGTTACCATTTCTGCTCTCTCTCCAAAGCCCT 60
450 ATGCCAGAAATAAAGAGATCTACCTGTTACCATTTCTGCTCTCTCTCCAAAGCCCT 509
61 GGAATGACAGACAGACAGTGCAGAAATGCTTACCTGATCCAGTCCAGTGCAGAGTGT 120
510 GGAATGACAGACAGACAGTGCAGAAATGCTTACCTGATCCAGTCCAGTGCAGAGTGT 569
121 TTGATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
570 TTGATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 629
```

```

: APPLICANT: Baker, Kevin P.
: APPLICANT: Chen, Jian
: APPLICANT: Desnoyers, Luc
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Pan, James
: APPLICANT: Smith, Victoria
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE OF INVENTION: ACIDS ENCODING THE SAME
: FILE REFERENCE: P3430R1C86
: CURRENT APPLICATION NUMBER: US/10/176,482
: PRIOR FILING DATE: 2002-06-20
: NUMBER OF SEQ ID NOS: 612
: LENGTH: 2609
: TYPE: DNA
: ORGANISM: Homo Sapien
US-10-176-482-407

```

```

Query Match      99.98; Score 1342.4; DB 9; Length 2609;
Best Local Similarity 99.98; Pred. No. 0;
Matches 1343; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 ATGCCGGAATAAAGGATATCTGCTTACCATTTCTGGCTCTCTGCTTCCAGCCCT 60
DB 450 ATGCCGGAATAAAGGATATCTGCTTACCATTTCTGGCTCTCTGCTTCCAGCCCT 509
QY 61 GGGATGACAGGACAGAGTGTGACGATGCTTGTGACGATGCTGACGATGCTGACGATGCT 120
DB 510 GGGATGACAGGACAGAGTGTGACGATGCTTGTGACGATGCTGACGATGCTGACGATGCT 569
QY 121 TTAGATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
DB 570 TTAGATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 629
QY 181 AACCAAAATGGGGGATTTATGATTTTATGATTTTATGATTTTATGATTTTATGATTTTATGAT 240
DB 630 AACCAAAATGGGGGATTTATGATTTTATGATTTTATGATTTTATGATTTTATGATTTTATGAT 689
QY 241 TCGAACCCCTACTGACAGCCCTACTGACAGCCCTACTGACAGCCCTACTGACAGCCCTACTGAC 300
DB 690 TCGAACCCCTACTGACAGCCCTACTGACAGCCCTACTGACAGCCCTACTGACAGCCCTACTGAC 749
QY 301 GCTCCAACTATCCAGAGATCTCCAGGCTCTTATATGCGCTTGTATGATGATGATGATGAT 360
DB 750 GCTCCAACTATCCAGAGATCTCCAGGCTCTTATATGCGCTTGTATGATGATGATGATGAT 809
QY 361 GAAAGAACCAATGTGTGATGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 810 GAAAGAACCAATGTGTGATGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 869
QY 421 ACCAGATCTGATCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 870 ACCAGATCTGATCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 929
QY 481 CTTCTGGAAGGCCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 930 CTTCTGGAAGGCCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 989
QY 541 TGTGGGAATGTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 990 TGTGGGAATGTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1049
QY 601 GATGGAAGGCTTTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 1050 GATGGAAGGCTTTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1109
QY 661 TGCCTCAACACCTACGCGCTTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGGAA 720

```

```

DB 1110 TGCCTCAACACCTACGCGCTTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGGAA 1169
QY 721 GATGCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 1170 GATGCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1229
QY 781 GATGCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 1230 GATGCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1289
QY 841 GATGACAAACGAGCTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 1290 GATGACAAACGAGCTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1349
QY 901 CTGACAGAGAGCTGCTCAATTTACAAAGGGGCTTCAATGATGATGATGATGATGATGATGATGAT 960
DB 1350 CTGACAGAGAGCTGCTCAATTTACAAAGGGGCTTCAATGATGATGATGATGATGATGATGATGAT 1409
QY 961 GAGGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 1410 GAGGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1469
QY 1021 TGCAGAGACAGCCCTTACCATCTGTACCGGGACATGAGAGTGTGACAGAGCTTCC 1080
DB 1470 TGCAGAGACAGCCCTTACCATCTGTACCGGGACATGAGAGTGTGACAGAGCTTCC 1529
QY 1081 GTTCCGCTGATCATCTTCCAAATGCAAGCCAGAGCCCTTACCTGCGGCTTATGATGAT 1140
DB 1530 GTTCCGCTGATCATCTTCCAAATGCAAGCCAGAGCCCTTACCTGCGGCTTATGATGAT 1589
QY 1141 TTCCAGATCAATCTGGGAATGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
DB 1590 TTCCAGATCAATCTGGGAATGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1649
QY 1201 AGTGCCACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
DB 1650 AGTGCCACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1709
QY 1261 GAATGATCACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 1710 GAATGATCACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1769
QY 1321 ATATATGTGTGCGAGTACCATTC 1344
DB 1770 ATATATGTGTGCGAGTACCATTC 1793

RESULT 13
US-10-176-757-407
: Sequence 407, Application US/10176757
: Publication No. US2003002297A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Chen, Jian
: APPLICANT: Desnoyers, Luc
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Pan, James
: APPLICANT: Smith, Victoria
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE OF INVENTION: ACIDS ENCODING THE SAME
: FILE REFERENCE: P3430R1C86
: CURRENT APPLICATION NUMBER: US/10/176,757
: PRIOR FILING DATE: 2002-06-20
: NUMBER OF SEQ ID NOS: 612
: LENGTH: 2609

```

```

Db      1590  TTCAGATCAAAATCTGGGAATGAGGGCAGAGAAATTTACATGGCGCAAAAGGGCCCATC 1649
QY      1201  AGTCCACACCTGGTGGATGATGACACAGCCCATCAAAAGGCCCGGGAAATCCAGCTGACTGG 1260
Db      1650  AGTCCACACCTGGTGGATGATGACACAGCCCATCAAAAGGCCCGGGAAATCCAGCTGACTGG 1709
QY      1261  GAAATGATCACTGTCACACACTGTGATCAACTTCAGAGGCAAGCTCCGATCGACTGGCGG 1320
Db      1710  GAAATGATCACTGTCACACACTGTGATCAACTTCAGAGGCAAGCTCCGATCGACTGGCGG 1769
QY      1321  ATATATGTGTGCGAGTACCATTTC 1344
Db      1770  ATATATGTGTGCGAGTACCATTTC 1793

RESULT 11
US-10-175-752-407
: Sequence 407, Application US/10175752
: Publication No. US20030022295A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Chen, Jian
: APPLICANT: Desnoyers, Luc
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Pan, James
: APPLICANT: Smith, Victoria
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: TITLE OF INVENTION: ACIDS ENCODING THE SAME
: FILE REFERENCE: P3430R1C60
: CURRENT APPLICATION NUMBER: US/10/175,752
: CURRENT FILING DATE: 2002-06-19
: Prior Application removed - See file wrapper or Palm
: NUMBER OF SEQ ID NOS: 612
: SEQ ID NO 407
: LENGTH: 2609
: TYPE: DNA
: ORGANISM: Homo Sapien
US-10-175-752-407

Query Match          99.9%; Score 1342.4; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1343; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  ATGCCAGGAATAAAGAGATCTCAGTCTGTTACATTCGATTCGGGCTCTGTCTTCAAGCCT 60
Db      450  ATGCCAGGAATAAAGAGATCTCAGTCTGTTACATTCGATTCGGGCTCTGTCTTCAAGCCT 509
QY      61  GGGAAATGCACAGGACAGTGCAGCAATGGCTTTGACCTGATTCGCCAGTCAGCAGTGT 120
Db      510  GGGAAATGCACAGGACAGTGCAGCAATGGCTTTGACCTGATTCGCCAGTCAGCAGTGT 569
QY      121  TTGATATTTGATGAAATCCCAACCATCCCGAGGCTTCGCGAGAGACATATGTGTGT 180
Db      570  TTGATATTTGATGAAATCCCGAACCATCCCGAGGCTTCGCGAGAGACATATGTGTGT 629
QY      181  AACCAAAATGGCGGGTATTTATGATTCGCCGGAACAAACCTGTGTATCGAGGGCCCTAC 240
Db      630  AACCAAAATGGCGGGTATTTATGATTCGCCGGAACAAACCTGTGTATCGAGGGCCCTAC 689
QY      241  TCGAACCCCTACTGACCCCCCTACTCAGGTCCGTACCCAGCAGCAGCTGCCCCACACTCTCA 300
Db      690  TCGAACCCCTACTGACCCCCCTACTCAGGTCCGTACCCAGCAGCAGCTGCCCCACACTCTCA 749
QY      301  GCTCCAAAGTATCCCAAGATCTCCAGGCGCTTTATATGCGGCTTTGATATACAGATGGAT 360
Db      750  GCTCCAAAGTATCCCAAGATCTCCAGGCGCTTTATATATGCGGCTTTGATATACAGATGGAT 809
QY      361  GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAACAGATTTCCACCAAGTCGAACCCC 420

```

Dd	810	GAAAGCAACCAATGTGTGGATGTGTGGACGAGTGTCCAAACAGTTTCCACACAGTGCACCCC	869
Oy	421	ACCAGATCTGCATCAATACTGGAAGCGGGGTACACCTGCTCTCTCACCGACGATATTGG	480
Dd	870	ACCAGATCTGCATCAATACTGGAAGCGGGGTACACCTGCTCTCTCACCGACGATATTGG	929
Oy	481	CTTGTGGAAGCCAGTGGTCTTGACATGTGAATGTGCGTATTGGTTACTGCGACAGCTC	540
Dd	930	CTTGTGGAAGCCAGTGGTCTTGACATGTGAATGTGCGTATTGGTTACTGCGACAGCTC	989
Oy	541	TGTGCGAATGTTCCTGGATTCCTATTCTTGTACATGCAACCCCTGGTATTTTACCCTCAATGAG	600
Dd	990	TGTGCGAATGTTCCTGGATTCCTATTCTTGTACATGCAACCCCTGGTATTTTACCCTCAATGAG	1049
Oy	601	GATGGAAAGTCTTTCGCAAGATGTGAACGAGTGTGCCACCCGAAACCCCTGGTGCMAACC	660
Dd	1050	GATGGAAAGTCTTTCGCAAGATGTGAACGAGTGTGCCACCCGAAACCCCTGGTGCMAACC	1109
Oy	661	TGCGTCAACACCTTACGCGCTTTTCAFCGCGCGCTGTGACCCAGATATGAACCTTGAGAA	720
Dd	1110	TGCGTCAACACCTTACGCGCTTTTCAFCGCGCGCTGTGACCCAGATATGAACCTTGAGAA	1169
Oy	721	GATGCGCTTCAATGTCAGATGATATGACAGAGTGCACGCTTCTGTGATTCCTCTGCCAAT	780
Dd	1170	GATGCGCTTCAATGTCAGATGATATGACAGAGTGCACGCTTCTGTGATTCCTCTGCCAAT	1229
Oy	781	GAGTGTGTGAACCAAGCCCGGACATATCTTTCGCTCGCCCTCCAGGCTACATCCTGCTG	840
Dd	1230	GAGTGTGTGAACCAAGCCCGGACATATCTTTCGCTCGCCCTCCAGGCTACATCCTGCTG	1289
Oy	841	GATGACAACCGAAGCTTGCACAAGACATCAACGAATGTGTAGCAGACAGAAACACACGCTGCAC	900
Dd	1290	GATGACAACCGAAGCTTGCACAAGACATCAACGAATGTGTAGCAGACAGAAACACACGCTGCAC	1349
Oy	901	CTGCAGACGACGCTCTCAATTTACAAGGGGGCTTCAATATGACATCGACCCCATCGCTGT	960
Dd	1350	CTGCAGACGACGCTCTCAATTTACAAGGGGGCTTCAATATGACATCGACCCCATCGCTGT	1409
Oy	961	GAGGAGCCTTATCTGAGAGATCAGTGAATTAACCGCTGTATGTGTCTCGTGTGAGAACCTGTGC	1020
Dd	1410	GAGGAGCCTTATCTGAGAGATCAGTGAATTAACCGCTGTATGTGTCTCGTGTGAGAACCTGTGC	1469
Oy	1021	TGCAGAGACACGCCCTTTACCATCTTGTACCGGGACATGACGTGTGTCAAGACGCTCC	1080
Dd	1470	TGCAGAGACACGCCCTTTACCATCTTGTACCGGGACATGACGTGTGTCAAGACGCTCC	1529
Oy	1081	GTTCGCCGCTGACATCTTCCAAATGCAAGCCACGACCCGCTACCTGGGGCCCTATTACATTT	1140
Dd	1530	GTTCGCCGCTGACATCTTCCAAATGCAAGCCACGACCCGCTACCTGGGGCCCTATTACATTT	1589
Oy	1141	TTCCAGATCAATCTGGGAATGAGGGCAGAGAAATTTTACATGTGGGCAAAACGGGCCCATC	1200
Dd	1590	TTCCAGATCAATCTGGGAATGAGGGCAGAGAAATTTTACATGTGGGCAAAACGGGCCCATC	1649
Oy	1201	AGTGGCACCCTGGTGATGATGACAGGCCCATCAAAAGGCCCGGGGAAATTCACCTGGACCTG	1260
Dd	1650	AGTGGCACCCTGGTGATGATGACAGGCCCATCAAAAGGCCCGGGGAAATTCACCTGGACCTG	1709
Oy	1261	GAAATGATCACTGTCAACACACTGTATCAAACTTCAGAGGACAGCTCCGTTATCCGACTGGG	1320
Dd	1710	GAAATGATCACTGTCAACACACTGTATCAAACTTCAGAGGACAGCTCCGTTATCCGACTGGG	1769
Oy	1321	ATATATGTGTGCGAGTACCATTC	1344
Dd	1770	ATATATGTGTGCGAGTACCATTC	1793

Db 1290 GATGACACCGAAGCTGCCAAGACATCAAGAAATGTGAGCAGAGAACCAACGACGTGAAC 1349
QY CTGCACACGACGCTGACATATTTACAGGGGGCTTCAATGATCGAGCCCATCCGCTGT 960
Db 1350 CTGCACACGACGCTGACATATTTACAGGGGGCTTCAATGATCGAGCCCATCCGCTGT 1409
QY GAGGACGCTTATCTGAGGATCAGATTAACCGCTGTATGTCTGCTGAGAACCTTGGC 1020
Db 1410 GAGGACGCTTATCTGAGGATCAGATTAACCGCTGTATGTCTGCTGAGAACCTTGGC 1469
QY 1021 TGCAGAGACCAAGCCCTTTTACCATCTTGTACCGGACATGACGTGTGTACGAGCTTC 1080
Db 1470 TGCAGAGACCAAGCCCTTTTACCATCTTGTACCGGACATGACGTGTGTACGAGCTTC 1529
QY 1081 GTTCCCGCTGACATCTTCCAAATGCAAGCCAGACCCGCTACCTTGGGGCTTATACAT 1140
Db 1530 GTTCCCGCTGACATCTTCCAAATGCAAGCCAGACCCGCTACCTTGGGGCTTATACAT 1589
QY 1141 TTCCAGATCAAAATCTGGGAATGAGGGCAGAAATTTTACATGCGGCAAGGGGCCCATC 1200
Db 1590 TTCCAGATCAAAATCTGGGAATGAGGGCAGAAATTTTACATGCGGCAAGGGGCCCATC 1649
QY 1201 AGTGCACCCCTGTGTATGACACGCCCATCAAGGGCCCGGAAATCCAGCTGACTTG 1260
Db 1650 AGTGCACCCCTGTGTATGACACGCCCATCAAGGGCCCGGAAATCCAGCTGACTTG 1709
QY 1261 GAAATGATCATGCTCAACACTGTATCATCACTTGTAGAGGACCTCGGTATCGAGCTGG 1320
Db 1710 GAAATGATCATGCTCAACACTGTATCATCACTTGTAGAGGACCTCGGTATCGAGCTGG 1769
QY 1321 ATATATGTGTGCGAGTACCATTC 1344
Db 1770 ATATATGTGTGCGAGTACCATTC 1793

RESULT 10
US-10-175-738-407
; Sequence 407, Application US/10175738
; Publication No. US2003022294A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jilan
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Collin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C45
; CURRENT APPLICATION NUMBER: US/10/175,738
; PRIORITY FILING DATE: 2002-06-19
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 407
; LENGTH: 2609
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-175-738-407

Query Match 99.98; Score 1342.4; DB 9; Length 2609;
Best Local Similarity 99.98; Pred. No. 0;
Matches 1343; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCCAGATATAAAGGATGATCTACATGTTACCATTTGGCTCTCTGTCTTCCAGAGCCT 60
Db 450 ATGCCAGATATAAAGGATGATCTACATGTTACCATTTGGCTCTCTGTCTTCCAGAGCCT 509
QY 61 GGGAAATGACAGGACAGTGCAGGAATGCTTGTACCTGATCGCAATGAGACATGT 120

Db 510 GGGAAATGACAGGACAGTGCAGGAATGCTTGTACCTGATCGCAATGAGACATGT 569
QY 121 TTGATATTTATGATGATCCGACACATCCCGAGGCTCCGAGAGACATGATGTGT 180
Db 570 TTGATATTTATGATGATCCGACACATCCCGAGGCTCCGAGAGACATGATGTGT 629
QY 181 AACCAAAATGGGGGATTTTATGATCCCGGAGCAAAACCTGTGTATGAGGGCTTC 240
Db 630 AACCAAAATGGGGGATTTTATGATCCCGGAGCAAAACCTGTGTATGAGGGCTTC 689
QY 241 TCGAACCCCTACTGACCCCTTACTAGTCCGTAACCGAGAGCTGCCCCACACTTCA 300
Db 690 TCGAACCCCTACTGACCCCTTACTAGTCCGTAACCGAGAGCTGCCCCACACTTCA 749
QY 301 GCTTCCAAATATCCCAAGATCTCCAGGCTCTTATATGCCCTTTGATACCATGAT 360
Db 750 GCTTCCAAATATCCCAAGATCTCCAGGCTCTTATATGCCCTTTGATACCATGAT 809
QY 361 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAAGATTCACACAGTGCACACCC 420
Db 810 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAAGATTCACACAGTGCACACCC 869
QY 421 ACCAGATCTGATCAATACTGAAGGGGGTACACCTGCTCTGACCGAGCATATGG 480
Db 870 ACCAGATCTGATCAATACTGAAGGGGGTACACCTGCTCTGACCGAGCATATGG 929
QY 481 CTTCGGAAGGCGAGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Db 930 CTTCGGAAGGCGAGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 989
QY 541 TGTGCAATGTCTGTGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Db 990 TGTGCAATGTCTGTGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1049
QY 601 GATGGAAGTCTTGCACAGATGTGACAGAGTGTGCCACCGAGAACCCCTGCTGCAACC 660
Db 1050 GATGGAAGTCTTGCACAGATGTGACAGAGTGTGCCACCGAGAACCCCTGCTGCAACC 1109
QY 661 TGGCTCAACACCTACAGGCTCTTATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 1110 TGGCTCAACACCTACAGGCTCTTATGATGATGATGATGATGATGATGATGATGATGAT 1169
QY 721 GATGCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 1170 GATGCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1229
QY 781 GAGTGTGTGAACAGCCCGGACATCTTGTCTCTGCTCCCTCAGGCTACATCTGCTG 840
Db 1230 GAGTGTGTGAACAGCCCGGACATCTTGTCTCTGCTCCCTCAGGCTACATCTGCTG 1289
QY 841 GATGACACCGGAAGCTCCCAAGACATCAACGAATGTGAGCACAGGAACACACGTCGAC 900
Db 1290 GATGACACCGGAAGCTCCCAAGACATCAACGAATGTGAGCACAGGAACACACGTCGAC 1349
QY 901 CTGCAGAGAGCTGCTCAATTTTACAGGGGGCTTCAATGATGATGATGATGATGATGAT 960
Db 1350 CTGCAGAGAGCTGCTCAATTTTACAGGGGGCTTCAATGATGATGATGATGATGATGAT 1409
QY 961 GAGGACGCTTATCTGAGGATCAGTATTAACCGCTGTATGTCTGCTGAGAACCTTGGC 1020
Db 1410 GAGGACGCTTATCTGAGGATCAGTATTAACCGCTGTATGTCTGCTGAGAACCTTGGC 1469
QY 1021 TGCAGAGACCAAGCCCTTTTACCATCTTGTACCGGACATGACGTGTGTACGAGCTTC 1080
Db 1470 TGCAGAGACCAAGCCCTTTTACCATCTTGTACCGGACATGACGTGTGTACGAGCTTC 1529
QY 1081 GTTCCCGCTGACATCTTCCAAATGCAAGCCAGACCCGCTACCTTGGGGCTTATACAT 1140
Db 1530 GTTCCCGCTGACATCTTCCAAATGCAAGCCAGACCCGCTACCTTGGGGCTTATACAT 1589
QY 1141 TTCCAGATCAAAATCTGGGAATGAGGGCAGAAATTTTACATGCGGCAAGGGGCCCATC 1200

```

DB 990 TGTGGAAATGTTCTTGATCTTATCTTGTATGACATGCAACCTGGTTTACCTCAANTAG 1049
OY 601 GATGGAAGCTTTGGCAAGATGTAAGAGATGTGCCACCGAGACCCCTGGTGTGAACC 660
DB 1050 GATGGAAGCTTTGGCAAGATGTAAGAGATGTGCCACCGAGACCCCTGGTGTGAACC 1109
OY 661 TGGCTCAACACCTAGAGCTCTTTCATCTGCGGCTGTGACCCAGATATGAACCTTGAGAA 720
DB 1110 TGGCTCAACACCTAGAGCTCTTTCATCTGCGGCTGTGACCCAGATATGAACCTTGAGAA 1169
OY 721 GATGCGCTTATGTCAGATGATATGAGACAGATGACAGCTTCTGTAGTTCTCTGCCACAT 780
DB 1170 GATGCGCTTATGTCAGATGATATGAGACAGATGACAGCTTCTGTAGTTCTCTGCCACAT 1229
OY 781 GAGTGTGTAAACACCGCCGACATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
DB 1230 GAGTGTGTAAACACCGCCGACATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1289
OY 841 GATGACAAACCGAAGCTGCGCAAGATCAACGAATGTGAGCAGACAGAACACACAGTGAAC 900
DB 1290 GATGACAAACCGAAGCTGCGCAAGATCAACGAATGTGAGCAGACAGAACACACAGTGAAC 1349
OY 901 CTGCAAGCAGAGCTGTCACAAATTTACAAAGGGGCTTCAAAATGCAATGCAACCCCTGCT 960
DB 1350 CTGCAAGCAGAGCTGTCACAAATTTACAAAGGGGCTTCAAAATGCAATGCAACCCCTGCT 1409
OY 961 GAGGAGCCTTATCTGAGATCACTGATTAACCGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 1410 GAGGAGCCTTATCTGAGATCACTGATTAACCGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 1469
OY 1021 TGCAGAGACACCCCTTTAACATCTTGTACCGGGACATGGAAGCTGTGTGAGAGAGCTTC 1080
DB 1470 TGCAGAGACACCCCTTTAACATCTTGTACCGGGACATGGAAGCTGTGTGAGAGAGCTTC 1529
OY 1081 GTTCCCGGTGACATCTTCAAAATGCAAGCAGACCGGCTAACCTGGGGCTTATTAAT 1140
DB 1530 GTTCCCGGTGACATCTTCAAAATGCAAGCAGACCGGCTAACCTGGGGCTTATTAAT 1589
OY 1141 TTCCAGATCAAAATCTGGAATGAGGGCAGAGAAATTTTACATGCGGCAAAACGGGCCATC 1200
DB 1590 TTCCAGATCAAAATCTGGAATGAGGGCAGAGAAATTTTACATGCGGCAAAACGGGCCATC 1649
OY 1201 AGTGCACCTCTGTGTGACACCCCTTAAAGGGGCTTCAAAATGCAATGCAACCCCTGCT 1260
DB 1650 AGTGCACCTCTGTGTGACACCCCTTAAAGGGGCTTCAAAATGCAATGCAACCCCTGCT 1709
OY 1261 GAAATGATCACTGTCAACACTGTCAATCACTTCAAGAGAGAGCTGCTGATCCGACTGGG 1320
DB 1710 GAAATGATCACTGTCAACACTGTCAATCACTTCAAGAGAGAGCTGCTGATCCGACTGGG 1769
OY 1321 ATATATGTGTGCGAGTACCATTC 1344
DB 1770 ATATATGTGTGCGAGTACCATTC 1793

```

RESULT 9
US-10-173-706-407
; Sequence 407, Application US/10173706
; Publication No. US2003002223A1

GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME

```

; FILE REFERENCE: P3430R1C7
; CURRENT APPLICATION NUMBER: US/10/173,706
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 407
; LENGTH: 2609
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-173-706-407

Query Match      99.9%; Score 1342.4; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1343; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ATGCCAGGAATTAAGATATCTACTGATTTACATTTGCGTCTGCTGCTTCCAGCCCT 60
DB 450 ATGCCAGGAATTAAGATATCTACTGATTTACATTTGCGTCTGCTGCTTCCAGCCCT 509
OY 61 GGGAAATGCACAGGACAGTGCACGAATGGCTTTGACCTGATGCGCAGTCAAGACAGT 120
DB 510 GGGAAATGCACAGGACAGTGCACGAATGGCTTTGACCTGATGCGCAGTCAAGACAGT 569
OY 121 TTAGATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
DB 570 TTAGATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 629
OY 181 AACCAAAATGGGGGTATTTATGATTTCCCGGCAAAACCTGTATGAGAGGCGCTTAC 240
DB 630 AACCAAAATGGGGGTATTTATGATTTCCCGGCAAAACCTGTATGAGAGGCGCTTAC 689
OY 241 TCGAACCCCTACTGACACCCCTTACTGAGTCCGTACCCAGACAGTGCAGCCCACTCTCA 300
DB 690 TCGAACCCCTACTGACACCCCTTACTGAGTCCGTACCCAGACAGTGCAGCCCACTCTCA 749
OY 301 GCTTCAAAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
DB 750 GCTTCAAAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 809
OY 361 GAAAGCAACCAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 810 GAAAGCAACCAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 869
OY 421 ACCAGATCTGATCAATTAAGAGGGGCTTACACCTCTCTGACACGAGGATATTTG 480
DB 870 ACCAGATCTGATCAATTAAGAGGGGCTTACACCTCTCTGACACGAGGATATTTG 929
OY 481 CTTCTGGAAGCCAGTCTTATGACATTTGATGATGATGATGATGATGATGATGATGAT 540
DB 930 CTTCTGGAAGCCAGTCTTATGACATTTGATGATGATGATGATGATGATGATGATGAT 989
OY 541 TGTGCGAATGTTCTGATCTTATCTTGTATGATGATGATGATGATGATGATGATGAT 600
DB 990 TGTGCGAATGTTCTGATCTTATCTTGTATGATGATGATGATGATGATGATGATGAT 1049
OY 601 GATGGAAGCTTTGGCAAGATGTAAGAGATGTGCCACCGAGAACCCCTGGTGTGAACC 660
DB 1050 GATGGAAGCTTTGGCAAGATGTAAGAGATGTGCCACCGAGAACCCCTGGTGTGAACC 1109
OY 661 TGGCTCAACACCTAGAGCTCTTTCATCTGCGGCTGTGACCCAGATATGAACCTTGAGAA 720
DB 1110 TGGCTCAACACCTAGAGCTCTTTCATCTGCGGCTGTGACCCAGATATGAACCTTGAGAA 1169
OY 721 GATGCGCTTATGTCAGATGATATGAGACAGATGACAGCTTCTGTAGTTCTCTGCCACAT 780
DB 1170 GATGCGCTTATGTCAGATGATATGAGACAGATGACAGCTTCTGTAGTTCTCTGCCACAT 1229
OY 781 GAGTGTGTAAACACCGCCGACATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
DB 1230 GAGTGTGTAAACACCGCCGACATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1289
OY 841 GATGACAAACCGAAGCTGCGCAAGATCAACGAATGTGAGCAGACAGAACACACAGTGAAC 900

```

Db 690 TCGAACCCCTACTGAGACCCCTACTGAGTCCGTACCCAGACGCTGCCACCACTCTCA 749
QY 301 GCTCCAACTATCCAGATCTCCAGCCTCTTATATGCGCCTTGTGATACAGATGAT 360
Db 750 GCTCCAACTATCCAGATCTCCAGCCTCTTATATGCGCCTTGTGATACAGATGAT 809
QY 361 GAAAGCAACCAATGTGGATGTGAGAGTGTGCAACAGATTTCCACCACTGCAACCC 420
Db 810 GAAAGCAACCAATGTGGATGTGAGAGTGTGCAACAGATTTCCACCACTGCAACCC 869
QY 421 ACCAGATCTGATCAATACTGAAAGCGGGGTACACTGCTCTCCACCGAGATATGG 480
Db 870 ACCAGATCTGATCAATACTGAAAGCGGGGTACACTGCTCTCCACCGAGATATGG 929
QY 481 CTTCGGAAGGCGAGTCTTAGACATTTGATGATGTGGCTATGCTACTGCGACAGCTC 540
Db 930 CTTCGGAAGGCGAGTCTTAGACATTTGATGATGTGGCTATGCTACTGCGACAGCTC 989
QY 541 TGTGCAATGTTCTCGATCTCTTATCTTGTATGATGCAACCCCTGTTTACCTCAATGAG 600
Db 990 TGTGCAATGTTCTCGATCTCTTATCTTGTATGATGCAACCCCTGTTTACCTCAATGAG 1049
QY 601 GATGGAAGTCTTGGCAAGATGTGAACGAGTGTCCACCGAAGCCCTGCGTCAACCC 660
Db 1050 GATGGAAGTCTTGGCAAGATGTGAACGAGTGTCCACCGAAGCCCTGCGTCAACCC 1109
QY 661 TCGCTCAACACCTAGAGGCTCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGAA 720
Db 1110 TCGCTCAACACCTAGAGGCTCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGAA 1169
QY 721 GATGCGTTCATTTGAGATATGAGACGAGTGTGCTCTGATGTTCTCTGCGCAACAT 780
Db 1170 GATGCGTTCATTTGAGATATGAGACGAGTGTGCTCTGATGTTCTCTGCGCAACAT 1229
QY 781 GAGTGTGAACACCGCGCGGCACTTCTCTGCTGCTGCGCTCCAGGCTACATCTGCTG 840
Db 1230 GAGTGTGAACACCGCGCGGCACTTCTCTGCTGCTGCGCTCCAGGCTACATCTGCTG 1289
QY 841 GATGCAACCGAAGTGTCCAGACATCAACGATGTGACACAGAACCCACAGCTGCAAC 900
Db 1290 GATGCAACCGAAGTGTCCAGACATCAACGATGTGACACAGAACCCACAGCTGCAAC 1349
QY 901 CTGAGCAGAGTGTGATCAATTTACAAAGGGGCTTCAAAATGCAATGCAACCCCATCCGCTG 960
Db 1350 CTGAGCAGAGTGTGATCAATTTACAAAGGGGCTTCAAAATGCAATGCAACCCCATCCGCTG 1409
QY 961 GAGAGCCTTATCTGAGATCAGTATACCGCTGATGTGCTGCTGCTGCTGCTGCTGCTG 1020
Db 1410 GAGAGCCTTATCTGAGATCAGTATACCGCTGATGTGCTGCTGCTGCTGCTGCTGCTG 1469
QY 1021 TGCAGACACGCGCTTACATCTTGTACCGGAGATGAGAGTGTGCTGAGAGCCTCC 1080
Db 1470 TGCAGACACGCGCTTACATCTTGTACCGGAGATGAGAGTGTGCTGAGAGCCTCC 1529
QY 1081 GTTCCCGCTGATCTTCAAAATGCAACGAGCGCGCTACCTGCGGCTTATTCATTT 1140
Db 1530 GTTCCCGCTGATCTTCAAAATGCAACGAGCGCGCTACCTGCGGCTTATTCATTT 1589
QY 1141 TTCCAGATCAAAATGTGGAAATGAGGCGAGAAATTTTACATGCGGCAAAAGGCGCCCATC 1200
Db 1590 TTCCAGATCAAAATGTGGAAATGAGGCGAGAAATTTTACATGCGGCAAAAGGCGCCCATC 1649
QY 1201 AGTGGCAGCCCTGGTATGAGACGCGCCCTCAAAAGGCGCGGGAATTCAGCTGAGCTTG 1260
Db 1650 AGTGGCAGCCCTGGTATGAGACGCGCCCTCAAAAGGCGCGGGAATTCAGCTGAGCTTG 1709
QY 1261 GAAATGATCACTGTCAACACTGTATCAACTTCAAGAGCAGTCTGCTGATGCGAGCTGG 1320
Db 1710 GAAATGATCACTGTCAACACTGTATCAACTTCAAGAGCAGTCTGCTGATGCGAGCTGG 1369
QY 1321 ATATATGTGTGCGAGTACCATTC 1344
Db 1770 ATATATGTGTGCGAGTACCATTC 1793

RESULT 8
US-10-175-737-407
Sequence 407, Application US/10175737
Publication No. US2003001315A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430RJC50
CURRENT APPLICATION NUMBER: US/10/175,737
PRIORITY FILING DATE: 2002-06-19
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 407
LENGTH: 2609
TYPE: DNA
ORGANISM: Homo Sapien
US-10-175-737-407

Query Match 99.9%; Score 1342.4; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;

Matches 1343; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCCAGAAATTAAGGATATCTCACTGTACCATTTCTGCTCTGCTTCCAGCCCT 60
Db 450 ATGCCAGAAATTAAGGATATCTCACTGTACCATTTCTGCTCTGCTTCCAGCCCT 509
QY 61 GGAATGACAGGACAGCACTGACAGATGAGCTTTGACCTGATGCGCAGTACAGAGCT 120
Db 510 GGAATGACAGGACAGCACTGACAGATGAGCTTTGACCTGATGCGCAGTACAGAGCT 569
QY 121 TTAGATATGATGATGATGCGCAACCATTCGCCAGGCTGCGGAGAGACATGATGTGT 180
Db 570 TTAGATATGATGATGATGCGCAACCATTCGCCAGGCTGCGGAGAGACATGATGTGT 629
QY 181 AACCAAAATGCGGGGATTTATGATTCCTCCGAGCAAAACCTGATGAGAGGCGCTAC 240
Db 630 AACCAAAATGCGGGGATTTATGATTCCTCCGAGCAAAACCTGATGAGAGGCGCTAC 689
QY 241 TCGAACCCCTTACTGAGCCCTTACTCAGTCCGTACCCAGAGAGTGCACCACTCTCA 300
Db 690 TCGAACCCCTTACTGAGCCCTTACTCAGTCCGTACCCAGAGAGTGCACCACTCTCA 749
QY 301 GCTCCAACTATCCAGATCTCCAGCCTCTTATATGCGCCTTGTGATACAGATGAT 360
Db 750 GCTCCAACTATCCAGATCTCCAGCCTCTTATATGCGCCTTGTGATACAGATGAT 809
QY 361 GAAAGCAACCAATGTGGATGTGAGAGTGTGCAACAGATTTCCACCACTGCAACCC 420
Db 810 GAAAGCAACCAATGTGGATGTGAGAGTGTGCAACAGATTTCCACCACTGCAACCC 869
QY 421 ACCAGATCTGATCAATACTGAAAGCGGGGTACACTGCTCTCCACCGAGATATGG 480
Db 870 ACCAGATCTGATCAATACTGAAAGCGGGGTACACTGCTCTCCACCGAGATATGG 929
QY 481 CTTCGGAAGGCGAGTCTTAGACATTTGATGATGTGGCTATGCTACTGCGACAGCTC 540
Db 930 CTTCGGAAGGCGAGTCTTAGACATTTGATGATGTGGCTATGCTACTGCGACAGCTC 989
QY 541 TGTGCAATGTTCTCGATCTCTTATCTTGTATGATGCAACCCCTGTTTACCTCAATGAG 600


```

; PRIOR APPLICATION NUMBER: PCT/US98/24855
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: PCT/US98/25108
; PRIOR FILING DATE: 1998-12-01
; PRIOR APPLICATION NUMBER: PCT/US98/25190
; PRIOR FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: PCT/US99/05028
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: PCT/US99/20111
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547

```

```

Query Match      99.9%; Score 1342.4; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1343; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 ATGCCGAGATTAAGAGATCTACTGTTACCATTTGCTCTCTCTTCCAAAGCCCT 60
DB 450 ATGCCGAGATTAAGAGATCTACTGTTACCATTTGCTCTCTCTTCCAAAGCCCT 509
QY 61 GGGAAATGACAGAGCAGAGTSCAGAAATGGCTTTGACCTGGATGGCCAGTCAGAGAGTGT 120
DB 510 GGGAAATGACAGAGCAGAGTSCAGAAATGGCTTTGACCTGGATGGCCAGTCAGAGAGTGT 569
QY 121 TTGATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
DB 570 TTGATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 629
QY 181 AACCAAAATGGCGGGATTTATGATTTCCCGGAGAAACCTGTGTATCGAGGGCCCTTAC 240
DB 630 AACCAAAATGGCGGGATTTATGATTTCCCGGAGAAACCTGTGTATCGAGGGCCCTTAC 689
QY 241 TCGAACCCCTACTCGAACCCCTACTCGAGTCCGTACCCAGAGTGGCCCAACACTCTCA 300
DB 690 TCGAACCCCTACTCGAACCCCTACTCGAGTCCGTACCCAGAGTGGCCCAACACTCTCA 749
QY 301 GCTCCAAATCTATCCCAAGATCTTCAGAGCTCTTATATGCGGCTTTGGATACCAAGATGAT 360
DB 750 GCTCCAAATCTATCCCAAGATCTTCAGAGCTCTTATATGCGGCTTTGGATACCAAGATGAT 809
QY 361 GAAACCAACCAATGTGTGATGTGAGAGTGTGCAACAGATTCGCCCAAGTGAACCC 420
DB 810 GAAACCAACCAATGTGTGATGTGAGAGTGTGCAACAGATTCGCCCAAGTGAACCC 869
QY 421 ACCGATCTGCATCAATATCTGAAAGCGGGTACACCTGTCTCTGCAACGAGATTTGG 480
DB 870 ACCGATCTGCATCAATATCTGAAAGCGGGTACACCTGTCTCTGCAACGAGATTTGG 929
QY 481 CTTCTGGAAGGCAATGCTTACACATTTGATGATGCTGCTATGCTTACTGCCAGCAGCTC 540
DB 930 CTTCTGGAAGGCAATGCTTACACATTTGATGATGCTGCTATGCTTACTGCCAGCAGCTC 989
QY 541 TGTGGAATGTTCTCGATACCTATCTTCTATACATGCAACCCGTTTACCCCAATGAG 600
DB 990 TGTGGAATGTTCTCGATACCTATCTTCTATACATGCAACCCGTTTACCCCAATGAG 1049
QY 601 GATGGAAGCTTGTGCAAGATGTGAACAGAGTGTGCCACCGAGAACCCCTGCTGCAAAC 660
DB 1050 GATGGAAGCTTGTGCAAGATGTGAACAGAGTGTGCCACCGAGAACCCCTGCTGCAAAC 1109
QY 661 TGGGCAACACCTAGCGGCTTTTCAATCTGCGGCTGTGACCCGAGATATGAATGAGAA 720
DB 1110 TGGGCAACACCTAGCGGCTTTTCAATCTGCGGCTGTGACCCGAGATATGAATGAGAA 1169
QY 721 GATGGCGTTCAATGAGTATGAGAGAGTGCAGCTTCTTGAGTTCCTGCTGCCAAACAT 780
DB 1170 GATGGCGTTCAATGAGTATGAGAGAGTGCAGCTTCTTGAGTTCCTGCTGCCAAACAT 1229

```

```

QY 781 GAGTGTGTGAACACGAGCCCGGACATCTTGTCTCTCTCCCTCCAGGCTACATCTGCTG 840
DB 1230 GAGTGTGTGAACACGAGCCCGGACATCTTGTCTCTCTCCCTCCAGGCTACATCTGCTG 1289
QY 841 GATGACAAACGGAAGCTGCGCAAGATCAACGAATGTGAGCAGACAGAACCAACGTCGAAC 900
DB 1290 GATGACAAACGGAAGCTGCGCAAGATCAACGAATGTGAGCAGACAGAACCAACGTCGAAC 1349
QY 901 CTGACAGACAGTGTCTCAATTTTCAAGGGGGCTTCAATGCAATGCAACCCATCCGCTGT 960
DB 1350 CTGACAGACAGTGTCTCAATTTTCAAGGGGGCTTCAATGCAATGCAACCCATCCGCTGT 1409
QY 961 GAGGACCTTATCTGAGATCAGTATACCGCTGTATGCTGCTGAGAACCCCTGCG 1020
DB 1410 GAGGACCTTATCTGAGATCAGTATACCGCTGTATGCTGCTGAGAACCCCTGCG 1469
QY 1021 TGCAGAGACAGCCCTTTTACATCTTGTATCCGGAGATGACAGCTGTGTGACAGAGCTTC 1080
DB 1470 TGCAGAGACAGCCCTTTTACATCTTGTATCCGGAGATGACAGCTGTGTGACAGAGCTTC 1529
QY 1081 GTTCCCGCTGACATCTTCCAAATGCAAGCCAGACCCGCTACCTGGGGCTTATTACAT 1140
DB 1530 GTTCCCGCTGACATCTTCCAAATGCAAGCCAGACCCGCTACCTGGGGCTTATTACAT 1589
QY 1141 TTCCAGATCAATCTGGAATGAGGGCAGAGAAATTTTACATGCGGCAACGCGCCCATC 1200
DB 1590 TTCCAGATCAATCTGGAATGAGGGCAGAGAAATTTTACATGCGGCAACGCGCCCATC 1649
QY 1201 AGTGCACCCCTGGTATGACAGCCCATTAAGGGGCGCGGGAATTCAGCTGAGCTTG 1260
DB 1650 AGTGCACCCCTGGTATGACAGCCCATTAAGGGGCGCGGGAATTCAGCTGAGCTTG 1709
QY 1261 GAAATGATCACTGTCAACACTGTATCAACTGAGAGGAGCTCCGTGATCCGACTGG 1320
DB 1710 GAAATGATCACTGTCAACACTGTATCAACTGAGAGGAGCTCCGTGATCCGACTGG 1769
QY 1321 ATATATGTTGCGAGTACCATTC 1344
DB 1770 ATATATGTTGCGAGTACCATTC 1793

```

RESULT 6 US-10-174-590-407

```

; Sequence 407, Application US/10174590
; Publication No. US20030008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C42
; CURRENT APPLICATION NUMBER: US/10-174,590
; PRIOR FILING DATE: 2002-06-18
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 407
; LENGTH: 2609
; TYPE: DNA
; ORGANISM: Homo sapien
; US-10-174-590-407

```

```

Query Match      99.9%; Score 1342.4; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;

```

PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/062285
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062816
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063082
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/063329
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063733
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066840
PRIOR FILING DATE: 1997-11-25
PRIOR APPLICATION NUMBER: 60/069694
PRIOR FILING DATE: 1997-12-16
PRIOR APPLICATION NUMBER: 60/074086
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/074092
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/095998
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/097000
PRIOR FILING DATE: 1998-08-18
PRIOR APPLICATION NUMBER: 60/099601
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099811
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100858
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101922
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/106032
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: 60/109304
PRIOR FILING DATE: 1998-11-20
PRIOR APPLICATION NUMBER: 60/125778
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/139655
PRIOR FILING DATE: 1999-06-15
PRIOR APPLICATION NUMBER: 60/145070
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/149396
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/169495
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 08/918874
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 08/933821
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 08/960507
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 09/114844
PRIOR FILING DATE: 1998-07-14
PRIOR APPLICATION NUMBER: 09/136801
PRIOR FILING DATE: 1998-08-19
PRIOR APPLICATION NUMBER: 09/136804

PRIOR FILING DATE: 1998-08-19
PRIOR APPLICATION NUMBER: 09/136828
PRIOR FILING DATE: 1998-08-19
PRIOR APPLICATION NUMBER: 09/158342
PRIOR FILING DATE: 1998-09-21
PRIOR APPLICATION NUMBER: 09/180997
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 09/202088
PRIOR FILING DATE: 1998-12-08
PRIOR APPLICATION NUMBER: 09/254311
PRIOR FILING DATE: 1999-03-03
PRIOR APPLICATION NUMBER: 09/254460
PRIOR FILING DATE: 1999-03-09
PRIOR APPLICATION NUMBER: 09/254465
PRIOR FILING DATE: 1999-03-05
PRIOR APPLICATION NUMBER: 09/284663
PRIOR FILING DATE: 1999-04-15
PRIOR APPLICATION NUMBER: 09/332928
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/332929
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/333075
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/333077
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380138
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380139
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/403296
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/403297
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/423741
PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: 09/423844
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 09/522342
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: 09/548815
PRIOR FILING DATE: 2000-04-13
PRIOR APPLICATION NUMBER: 09/664610
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: 09/665350
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: 09/709238
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 09/767609
PRIOR FILING DATE: 2001-01-22
PRIOR APPLICATION NUMBER: 09/802706
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 09/808689
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 09/866028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 09/870574
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: 09/872035
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 09/886342
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: PCT/US98/14552
PRIOR FILING DATE: 1998-07-14
PRIOR APPLICATION NUMBER: PCT/US98/18824
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: PCT/US98/19093
PRIOR FILING DATE: 1998-09-14
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: PCT/US98/19437
PRIOR FILING DATE: 1998-09-17

SEQUENCE CHARACTERISTICS:
 LENGTH: 2362 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-10-041-016-1

Query Match 100.0%; Score 1344; DB 9; Length 2362;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGCCGAGATTAAGAGATCTACATCTGTTACCATCTGCGCTCTGCTCTCCAAAGCCCT 60
DB 186 ATGCCGAGATTAAGAGATCTACATCTGTTACCATCTGCGCTCTGCTCTCCAAAGCCCT 245
QY 61 GGGATGACACAGGCACAGTGCAGCAATGGCTTTGACCTGATCGCCAGTCAAGACAGTGT 120
DB 246 GGGATGACACAGGCACAGTGCAGCAATGGCTTTGACCTGATCGCCAGTCAAGACAGTGT 305
QY 121 TTAGATATTGATTAATGCGGCAACCATCCCGAGGCTGCGGAGAGACATGATGTGT 180
DB 306 TTAGATATTGATTAATGCGGCAACCATCCCGAGGCTGCGGAGAGACATGATGTGT 365
QY 181 AACCAAAATGGCGGGTATTATGATCCCGGACAAACCTGTGTATCCAGGGCCCTAC 240
DB 366 AACCAAAATGGCGGGTATTATGATCCCGGACAAACCTGTGTATCCAGGGCCCTAC 425
QY 241 TCGAACCCCTACTGACACCCCTACTAGATCCGTACCCAGAGAGCTGCCACCACTCTCA 300
DB 426 TCGAACCCCTACTGACACCCCTACTAGATCCGTACCCAGAGAGCTGCCACCACTCTCA 485
QY 301 GCTCCAAATATCCCAAGATCTCCAGGCTCTTATATGCGGCTTTGGATACCAATGAT 360
DB 486 GCTCCAAATATCCCAAGATCTCCAGGCTCTTATATGCGGCTTTGGATACCAATGAT 545
QY 361 GAAGCAACCAATGTGTGATGTGACAGAGTGTGCAACAGATTCCCAAGTGAACCC 420
DB 546 GAAGCAACCAATGTGTGATGTGACAGAGTGTGCAACAGATTCCCAAGTGAACCC 605
QY 421 ACCGAGATGTGATCAATCTGTAAGGCGGGGTACACCTGCTCTGACCGAGAGATATGG 480
DB 606 ACCGAGATGTGATCAATCTGTAAGGCGGGGTACACCTGCTCTGACCGAGAGATATGG 665
QY 481 CTTTGGAAAGGCGAGGCTTAGACATGTGATGTGCGTATGCTTACTGCGACAGCTC 540
DB 666 CTTTGGAAAGGCGAGGCTTAGACATGTGATGTGCGTATGCTTACTGCGACAGCTC 725
QY 541 TGTGGGAATGTTCTGTGATCTTATCTGTACATGCAACCTGGTTTACCTCAATGAG 600
DB 726 TGTGGGAATGTTCTGTGATCTTATCTGTACATGCAACCTGGTTTACCTCAATGAG 785
QY 601 GATGAAGGCTCTTGCAGATGTGAACGAGTGTGCCACGAGAACCCCTGCGCAAC 660
DB 786 GATGAAGGCTCTTGCAGATGTGAACGAGTGTGCCACGAGAACCCCTGCGCAAC 845
QY 661 TGGCGTCAACCTAGGCGCTTTTCATCTGCGCTGTGACCCAGAGATGAACTTGAGAA 720
DB 846 TGGCGTCAACCTAGGCGCTTTTCATCTGCGCTGTGACCCAGAGATGAACTTGAGAA 905
QY 721 GATGGCGTTTCATCTGAGTATGAGACAGTCAAGTCTCTGAGTCTCTGCGCAACAT 780
DB 906 GATGGCGTTTCATCTGAGTATGAGACAGTCAAGTCTCTGAGTCTCTGCGCAACAT 965
QY 781 GAGTGTGGAACACAGCCCGGACATCTCTGCTCTGCTGCGCAAGCATCTGCTG 840
DB 966 GAGTGTGGAACACAGCCCGGACATCTCTGCTCTGCTGCGCAAGCATCTGCTG 1025
QY 841 GATGACAACGAGAGTGCACAGATCAACGAATGTGACACAGAAACACAGTGTCAAC 900
DB 1026 GATGACAACGAGAGTGCACAGATCAACGAATGTGACACAGAAACACAGTGTCAAC 1085

```

```

QY 901 CTGCAGCAGACGTGCTACAAATTTCAAGGGGGCTTCAATGCAATGCCACCCGCTGT 960
DB 1086 CTGCAGCAGACGTGCTACAAATTTCAAGGGGGCTTCAATGCAATGCCACCCGCTGT 1145
QY 961 GAGGAGCCTTATCTGAGGATCAGTGAATPACCCGTGTATGTCTCTGTGAGAACCTGGC 1020
DB 1146 GAGGAGCCTTATCTGAGGATCAGTGAATPACCCGTGTATGTCTCTGTGAGAACCTGGC 1205
QY 1021 TGCAGAGACAGCCCTTACCATCTGTACCGGGACATGAGACGTGTGTGAGACGCTCC 1080
DB 1206 TGCAGAGACAGCCCTTACCATCTGTGTACCGGGACATGAGACGTGTGTGAGACGCTCC 1265
QY 1081 GTTCCGCTGACATCTTCCAAATGCAAGCCAGCAGCCGCTACCTGGGGCTTATACAT 1140
DB 1266 GTTCCGCTGACATCTTCCAAATGCAAGCCAGCAGCCGCTACCTGGGGCTTATACAT 1325
QY 1141 TTCCAGATCAAACTGGGATGAGAGGAGAGAAATTTTACATGCGGCAAAAGCCGCCATC 1200
DB 1326 TTCCAGATCAAACTGGGATGAGAGGAGAGAAATTTTACATGCGGCAAAAGCCGCCATC 1385
QY 1201 AGTCCACCCCTGTGTATGACACGCCCATCAAGGGCCCGGAAATTCAGCTGAGCTTG 1260
DB 1386 AGTCCACCCCTGTGTATGACACGCCCATCAAGGGCCCGGAAATTCAGCTGAGCTTG 1445
QY 1261 GAATATGATCACTGTCAACACTGTCACTTCACTGAGAGGAGCTCGTATCCGACTGGG 1320
DB 1446 GAATATGATCACTGTCAACACTGTCACTTCACTGAGAGGAGCTCGTATCCGACTGGG 1505
QY 1321 ATATATGTGTGAGTACCATTC 1344
DB 1506 ATATATGTGTGAGTACCATTC 1529

```

RESULT 3

US-09-083-002-1

Sequence 1, Application US/09083002

Patent No. US2001001650A1

GENERAL INFORMATION:

```

APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Racie, Lisa A.
APPLICANT: Lavallee, Edward R.
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Evans, Cheryl
APPLICANT: Agostino, Michael
APPLICANT: Lu, Zhijian
APPLICANT: Honjo, Tasuku
APPLICANT: Tashiro, Kei
APPLICANT: Nakamura, Tomoyuki
TITLE OF INVENTION: SECRETED PROTEINS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083,002
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: P-41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284

```

MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 211..1554
 US-09-275-805-1

Query Match 100.0%; Score 1344; DB 10; Length 1717;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGCCGGAATATAAAGATACACCTGTTACCATTTCTGGCTCTCTGCTTCCAAAGCCCT 60
DB 211 ATGCCGGAATATAAAGATACACCTGTTACCATTTCTGGCTCTCTGCTTCCAAAGCCCT 270
QY 61 GGAATGACAGCAGCAGTGTGATGCTTTGACCGCGATCGGACGACGAGTGT 120
DB 271 GGAATGACAGCAGCAGTGTGATGCTTTGACCGCGATCGGACGACGAGTGT 330
QY 121 TTGATATTTGATGATATGCGAACCATCCCGAGGCGCTGCGAGAGACATGATGTGT 180
DB 331 TTGATATTTGATGATATGCGAACCATCCCGAGGCGCTGCGAGAGACATGATGTGT 390
QY 181 AACCAAAAGGGGGGATTTATGATTCGCCGGAACAAACCGTGTATCGAGGCGCTAC 240
DB 391 AACCAAAAGGGGGGATTTATGATTCGCCGGAACAAACCGTGTATCGAGGCGCTAC 450
QY 241 TCGAACCCCTACTGACACCCCTACTAGTCCGCTACCCAGCAGCTGCCCACTCTCA 300
DB 451 TCGAACCCCTACTGACACCCCTACTAGTCCGCTACCCAGCAGCTGCCCACTCTCA 510
QY 301 GCTCCAACTATCCACGATCTCCAGGCTCTCTTATATGCGGCTTTGATACAGATGAT 360
DB 511 GCTCCAACTATCCACGATCTCCAGGCTCTCTTATATGCGGCTTTGATACAGATGAT 570
QY 361 GAAAGAACCAATGTGTGATGTGAGAGATGTGCAACAGATTCACACAGTGCACACCC 420
DB 571 GAAAGAACCAATGTGTGATGTGAGAGATGTGCAACAGATTCACACAGTGCACACCC 630
QY 421 ACCCAGATCTGACATACATGTAAGGCGGTACACCTGCTCTGACAGCGAGATTTGG 480
DB 631 ACCCAGATCTGACATACATGTAAGGCGGTACACCTGCTCTGACAGCGAGATTTGG 690
QY 481 CTTCTGGAAGGCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 540
DB 691 CTTCTGGAAGGCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 750
QY 541 TGTGCAATGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
DB 751 TGTGCAATGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 810
QY 601 GATGGAAGTCTTGGCAAGATGTGAAGAGTGTGCAACCGGACCCCTGCTGCAAAAC 660
DB 811 GATGGAAGTCTTGGCAAGATGTGAAGAGTGTGCAACCGGACCCCTGCTGCAAAAC 870
QY 661 TGGGTCAACACCTACGCGCTCTTCACTGCGGCTGTGACCCAGAGATATGAACCTG 720
DB 871 TGGGTCAACACCTACGCGCTCTTCACTGCGGCTGTGACCCAGAGATATGAACCTG 930
QY 721 GATGGGCTTCACTGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 780
DB 931 GATGGGCTTCACTGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 990
QY 781 GAGTGTGTGAACAGCCCGGACACATCTTCTGCTCTGCTCTGCTCTGCTCTGCTG 840
DB 991 GAGTGTGTGAACAGCCCGGACACATCTTCTGCTCTGCTCTGCTCTGCTCTGCTG 1050
QY 841 GATGACACCGAAGCTGCAAGATGATGATGATGATGATGATGATGATGATGATGATG 900
DB 1051 GATGACACCGAAGCTGCAAGATGATGATGATGATGATGATGATGATGATGATGATG 1110
QY 901 CTGCGACGAGCGGTCAATTTTCAAGGGGGCTTCAATGATGATGATGATGATGATG 960
DB 1111 CTGCGACGAGCGGTCAATTTTCAAGGGGGCTTCAATGATGATGATGATGATGATG 1170

```

```

QY 961 GAGAGCCTTATCTGAGGATCAGTATPAACCGCTGTATGTCTGCTGAGAACCCTGCG 1020
DB 1171 GAGAGCCTTATCTGAGGATCAGTATPAACCGCTGTATGTCTGCTGAGAACCCTGCG 1230
QY 1021 TGCAGAGACCAACCCCTTTACCATCTTGTACCGGACATGACGATGCTGACAGCCTCC 1080
DB 1231 TGCAGAGACCAACCCCTTTACCATCTTGTACCGGACATGACGATGCTGACAGCCTCC 1290
QY 1081 GTTCCCGCTGACATCTTCCAAATGCAAGCCACGACCCGCTACCTGCGGCTTATTAATT 1140
DB 1291 GTTCCCGCTGACATCTTCCAAATGCAAGCCACGACCCGCTACCTGCGGCTTATTAATT 1350
QY 1141 TTCCAGATCAAAATGTGGGAATGAGGGGAGAGAAATTTTACATGCGGCAACGGGCCCATC 1200
DB 1351 TTCCAGATCAAAATGTGGGAATGAGGGGAGAGAAATTTTACATGCGGCAACGGGCCCATC 1410
QY 1201 AGTGCACCCCTGTGTATGACACGCCCCATCAAAAGGGCCCCGGAAATCCAGCTGACTTG 1260
DB 1411 AGTGCACCCCTGTGTATGACACGCCCCATCAAAAGGGCCCCGGAAATCCAGCTGACTTG 1470
QY 1261 GAAATGATCACTGTCACACACTGTATCAACTTCAGAGGACGCTCGTGTATCCGACTGCG 1320
DB 1471 GAAATGATCACTGTCACACACTGTATCAACTTCAGAGGACGCTCGTGTATCCGACTGCG 1530
QY 1321 ATATATGTGTGCGAGTACCCATTTC 1344
DB 1531 ATATATGTGTGCGAGTACCCATTTC 1554

```

RESULT 2 US-10-041-016-1

Sequence 1, Application US/10041016
 Patent No. US20020165151A1

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth

McCoey, John M.

Racie, Lisa A.

Lavallie, Edward R.

Merberg, David

Treacy, Maurice

Evans, Cheryl

Agostino, Michael

Lu, Zhijian

Honjo, Tasuku

TITLE OF INVENTION: SECRETED PROTEINS

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genetics Institute, Inc.

STREET: 87 Cambridgepark Drive

CITY: Cambridge

STATE: MA

COUNTRY: U.S.A.

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/041,016

FILING DATE: 07-Jan-2002

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/083,002

FILING DATE: 21-Mar-1998

ATTORNEY/AGENT INFORMATION:

NAME: Sprunger, Suzanne A.

REGISTRATION NUMBER: P-41,323

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 498-8284

TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 1:

STRANDEDNESS: single

TOPOLOGY: Linear

Sat Jul 5 15:31:05 2003

us-09-674-379a-11.rge

Page 16

Oy 1297 GGCAGCTCCGTATCCGACTGGCGATATATGTCGCACTAACCATTC 1344
| | | | | | | | | | | | | |
Db 1378 GCCAGCTCTGTACGAGGCTCACCGTCTTTGAAGGGCTACACCCTTC 1425

Search completed: July 3, 2003, 15:34:23
Job time : 3509.57 secs

QY		721	GATGGCGCTTCATTGTCAGTGTATATGGACGAGTGCGAGTTCTCTGAGTTCCCTGCCCAACAT	780
Db		1040	GATGGCATTTCACGTGCATCATGTANTATGGACGAGTGCGAGTTCTCTGAGTTCCCTGCCCAAC	1099
OY		781.	GAGTGTGTGAACCAAGCCCCGACACTTCTTGCTCCGCCCCTCCAGGCTACATCTGCTG	840
Db		1100	GAGTGTGTGAACCAAGCCCCGCTCATCTTCTGCTGCGCCCTCCAGGCTACGCTGCTG	1159
OY		841	GATGACAACCGAAGCTGCCAAAGATCATAGAAGTAATGATGACACAGAACCCACAGTGCAC	900
Db		1160	GATGATAACCGAAGCTGCCAGGATATCATTAATGATGACACCGAACCACACGTATACC	1219
OY		901	CTGCAGCAGACGTGCTCAAAATTTTCAAGGGGGCTTCCAATGATCATGACCCCATCGCGT	960
Db		1220	TCACTGCAGACTTCTCAAAATCTCAAGGGGGCTTCCAATGATGATTCATGATCCATGCTGT	1279
OY		961	GAGGAGCCTTATCTGAGSAGTACGATATACCGCTGTATGTGTCTCTGCTGAGACCCCTGGC	1020
Db		1280	GAGGAGCCTTATCTGCTGATGTGGTGAACACCGCTGTATGTGTCTCTGCTGAGACACCAAGC	1339
OY		1021	TGCAGAGACCAAGCCCTTTCACATTTTGATCCGGGACATGGAACGTTGGTGCAGACGGCTCC	1080
Db		1340	TGCAGAACCAAGCCATTCACCATCTCTGATGCGSACATGATGTGGTGTGTCAGGACGCTCC	1399
OY		1081	GTTCCCGCTGACATCTTCCAAATGCAAGCACGACCCGCTACCTGTGGGGCTTATTCATT	1140
Db		1400	GTTCCCTCTGACATCTTCCAGATTCAGAACCAACCCGATACCTGTGGCTTATTCATT	1459
OY		1141	TTCCAGATCAAAATGTGGGAATGAGGGGAGAAATTTATACATGGGCAAAGGGGCCCATC	1200
Db		1460	TTCCAGATCAAAATGTGGCAACGAGGGTGCAGAGTTCTATATGGCGCAACAGGGGCTATTC	1519
OY		1201	AGTGCACACCTGTGTGATGACACGCCCCCATCAAGGGGCCCGGGAAATCCAGCTGGACTTG	1260
Db		1520	AGTGCACACCTGTGTGATGACACGCCCCCATCAAGGGGCTCGGGACATCCAGCTGGACTTG	1579
OY		1261	GAATGATTCACCTGTCACACACTGTATATCAACTTCAGAGGACAGCTCCGATCCGACTCGG	1320
Db		1580	GAGTGTGATTCACCTGTCACACACTGTATATCAACTTCAGAGGACAGCTCCGATCCGACTCGG	1639
OY		1321	ATATATGTGTGCGAGTACCCATTG	1344
Db		1640	ATATATGTGTGCGAGTATCCGTTG	1663
RESULT 12				
AFI12153				
LOCUS		2499 bp	mRNA	linear
DEFINITION		Rattus norvegicus developmental arteries and neural crest EGF-like protein mRNA, complete cds.		ROD 02-AUG-1999
ACCESSION		AFI12153		
VERSION		AFI12153.1	GI:5305674	
KEYWORDS				
SOURCE		Rattus norvegicus.		
ORGANISM		Rattus norvegicus		
REFERENCE				
AUTHORS		Nakamura,T., Ruiz-Jozano,P., Lindner,V., Yabe,D., Taniwaki,M., Furukawa,K., Kobuke,K., Tashiro,K., Iu,Z., Andon,N.L., Schaub,R., Matsumori,A., Sasayama,S., Chien,K.R. and Honjo,T.		
TITLE		DANCE, a novel secreted RGD protein expressed in developing, atherosclerotic, and balloon-injured arteries		
JOURNAL		J. Biol. Chem. 274 (32), 22476-22483 (1999)		
MEDLINE		99357779		
PUBMED		10428823		
REFERENCE		2 (bases 1 to 2499)		
AUTHORS		Nakamura,T., Yabe,D., Tashiro,K. and Honjo,T.		
TITLE		Direct Submission		
JOURNAL		Submitted (09-DEC-1998) Medicine, University of California San Diego,9500 Gilman Drive, La Jolla, CA 92093-0613, USA		

FEATURES	source	Location/Qualifiers
CDS		1..2499 /organism="Rattus norvegicus" /db_xref="taxon:10116" 350..1696 /note="DANCE; contains 6 cBEGF domains and 1 RGD motif; secreted protein; expressed in developmental aorta, neural crest cells, balloon injured vessels, atherosclerotic lesions" /codon_start=1 /product="developmental arteries and neural crest EGF-like protein" /protein_id="AAD1769.1" /db_xref="GI:5305675" /translation="MPGLRLITVTILALPHRPNAAQQCTNGFDLDRQGLDIDIECPRIAPACRGDMKCVNONGVLCIPRNPVRYRPSNYSSTSGVPAAPAPVPSNYPIISRLPYCRREYQMDENQCYDVECATDSHOCNPQICINTEGGYTCSDITGWLTBQOQJIDIDECRYGCOQLANVPYSICTCNQGFILNDGRCDVNEETNPVQVCNNTYGSDFICRCDDPYELEEDGIIHSDMDGCSFSELCQHEKNVDPGSYFCSCPRGVYLLEDRNSCODINECHRNHTCTPLQVCMYAGCFICIDPIVCEEPYLILDNRMCMQPAENLTGRDOEFTLLFRDMIVYPSRSVPADIFQOATTRYGAVYIPIQIKSGNREGEYMRQGISATLVMTRPILKGPRLDLDLEMITVMTVINFRRSSVILRIYVSQPF"
BASE COUNT	584 a 669 c 595 g 649 t	2 others
ORIGIN		
Query Match	82.1%; Score 1104; DB 10; Length 2499;	
Best Local Similarity	88.8%; Pred. No. 5e-309;	
Matches 1194; Conservative 0; Mismatches 150; Indels 0; Gaps 0;		
OY	1 ATGCAGAAATAAAGATACACAGCTTACCATTCGCGCTCTGCTTCCAAACCCCT 60	
Db	350 ATGCAGAGATTAAAGATACCTACCTGTCACCATCTCGGACAGCTGGCTTCCACATCT 409	
OY	61 GGGATGACACAGGCACAGTGCACGAATGGCTTTGACCTGGATCCCACTCAGACAGTGT 120	
Db	410 GGGATGACACAGCAACAGTGCACCAAGCGCTTTGACCTGGACCGCAGACAGACAGAGT 469	
OY	121 TTAATATATTGATGATGCCAACCATCCCGAGCGCTCGCAGAGACATATGTGTGTT 180	
Db	470 TTAATATATTGATGATGCGGACCATCCCTGAGCGCTTGGCGGACATGATGTGTGTC 529	
OY	181 AACCAAAATGCGGGGTATTATGATATCCCGGCAAAACCTGTGATCGAGGGCCCTAC 240	
Db	530 AACCAAGATGCGGGGTATCTGTGATCCCTGGAACCAACCAAGTATCGAGGGCCCTTAC 589	
OY	241 TCGAACCCCTACTGCAGCCCTACTCAGGTCCGTACCCAGCAGCTGCCCCACACTCTCA 300	
Db	590 TCCAAATCCCTACTACTACTCTACTCAGGCCCAATACCAGCAGCGCCACCACTAGTCCA 649	
OY	301 GCTCAAACTTCCACAGATCTCCAGGCGCTTATATAGCCGCTTTGGATTACCAATGGAT 360	
Db	650 GCTTCAACTTACCCACAGATTCCAGGCGCTTGTGTCTGTGCGCTTGGGTATCAATGGAT 709	
OY	361 GAAAGCAACCAATGTGTGATGTGGAGAGTGTGCACAGATTTCCACAGTGCACACCCC 420	
Db	710 GAAAGCAACCAAGTGTGGATGTGGAGAGTGTGCACAGATTTCCACACAGTGCACACCC 769	
OY	421 ACCAGATCTGCATCAATACTGAAAGGGGGGTACACACTGCTCTCTGCACCGAGCATATTGG 480	
Db	770 ACCAGATCTGTATCAACACGGAAGGAGGTACACCTGCTCTCAGCATGAGGGGTACTGG 829	
OY	481 CTTTGTGAAGCCAGTGTGACATTTGATTAATGTGCTATGGTTACTGTGCACAGCTC 540	
Db	830 CTTTGTGAAGGGCAGTGTGATTAATGTGATTAATGTGCTATGGTTACTGTGCACAGCTC 889	
OY	541 TGTGCGAATGTCTCGATCTCTATCTTTGATACATGCAACCCCTGTGTTTACCCTCAATGAG 600	
Db	890 TGTGCGAATGTCTCGATCTCTATCTTTGATGCTGTAACCCCTGTGCTTCAACGAT 949	
OY	601 GATGGAAGGCTTGTGCCAAGATGTGAAGAGTGTGCCACCGAAGACCCCTGGTGTGCAAAAC 660	
Db	950 GATGGAAGGCTTGTGCCAAGATGTGAAGAGTGTGAACCTGGAACCCCTGTGTTCAAGACC 1009	

Db 780 GATGGCATTCACGTCAGTGTATGAGCAGATGCGCTTCTCGAGTTCTTGTCACAC 839
Qy 781 GAGGTGTGAACACGCGCGGACATCTTCTGCTCTGCTCCAGGCTACATCTGCTG 840
Db 840 GAGTGTGTAAACGCGCGGCTCATCTTCTGCTGCTGCTCCAGGCTACGCTGTTG 899
Qy 841 GATGACACCGGAGCGCGGACATCAACGAATGTAGCAGCAGAACACACGTCGAC 900
Db 900 GATGATACCGGAAGCTGCGCAGGATATCAAGAAATGAGCAGCAGAACACACGTCG 959
Qy 901 CTGACAGACAGCTGTCTACATTTTACAAAGGGGGCTTCAATATGATGAGCCCGCTG 960
Db 960 TCACGTCGACACTGTGTACATCTACAGAGGGGCTTCAAAATGTATGATCCATGCTGT 1019
Qy 961 GAGGAGCCTTATCTGAGATCACTGATACACCGCTGATGTCTGCTGAGAACCTGCG 1020
Db 1020 GAGGAGCCTTATCTGCTGATGTGAAACCGCTGATGTCTGCTGAGCAGACAGC 1079
Qy 1021 TGCAGAGACAGCGCTTACCATCTGTACCGGAGACGTCGTCGTCAGACGCTCC 1080
Db 1080 TGCAGAGACAGCGCTTACCATCTGTATCGGAGATGATGTGTCGAGACGCTCC 1139
Qy 1081 GTTCCCGCTGACATCTTCCAAATGACAGCAGCAGCAGCTACCTGCGGCTTATTAAT 1140
Db 1140 GTTCCCGCTGACATCTTCCAGATGACAGAACACCGGATACCTGCTGCTATTAAT 1199
Qy 1141 TTCGAGATCAATCTGAGATGAGGAGCAGAAATTTTACATGCGGCGAAACGCGCCATC 1200
Db 1200 TTCGAGATCAATCTGAGCAGAGGAGGTCGAGATTTATATGCGGCGAAACGCGCTATC 1259
Qy 1201 AGTGCACACCTGTGTATGACAGCGCCCATCAAGGGGCGCGGAAATCCAGCTGACTG 1260
Db 1260 AGTGCACACCTGTGTATGACAGCGCCCATCAAGGGGCGCTGCGACATCAAGCTGACTG 1319
Qy 1261 GAAATGATCACTGTCAACACTGTCACTTCACTGAGAGGAGCTCCGTCATCCGACTGCG 1320
Db 1320 GAGATGATCACTGTCAACACTGTCACTTCACTGAGAGGAGCTCCGTCATCCGACTGCG 1379
Qy 1321 ATATATGTGTGCGAGTACCATTC 1344
Db 1380 ATATATGTGTGCGAGTATCCGTTT 1403

RESULT 11
AF112151 2478 bp mRNA linear ROD 02-AUG-1999
LOCUS AF112151
DEFINITION Mus musculus developmental arteries and neural crest EGF-like
ACCESSION AF112151.1 GI:5305670
VERSION
KEYWORDS
SOURCE
ORGANISM Mus musculus.
Eutheria: Chordata: Craniata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Mus:
1 (bases 1 to 2478)
Nakamura, T., Ruiz-Lozano, P., Lindner, V., Yabe, D., Tanisaki, M.,
Furukawa, Y., Kohike, K., Tashiro, K., Lu, Z., Andon, N.L., Schaub, R.,
Matsumori, A., Sasayama, S., Chien, K.R. and Honjo, T.
DANCE, a novel secreted RGD protein expressed in developing,
atherosclerotic, and balloon-injured arteries
J. Biol. Chem. 274 (32), 22476-22483 (1999)
JOURNAL MEDLINE
PUBMED 10428823
REFERENCE 2 (bases 1 to 2478)
AUTHORS Nakamura, T., Yabe, D., Tashiro, K. and Honjo, T.
TITLE Direct Submission
JOURNAL Submitted (09-DEC-1998) Medicine, University of California San
Diego, 9500 Gilman Drive, La Jolla, CA 92093-0613, USA
FEATURES
source
1..2478
/organism="Mus musculus"

CDS
/db_xref="taxon:10090"
320..1666
/note="DANCE; contains 6 cDEG domains and 1 RGD motif;
secreted protein; expressed in developmental aorta, neural
crest cells, balloon injured vessels, atherosclerotic
lesions"
/codon_start=1
/product="developmental arteries and neural crest EGF-like
protein"
/db_xref="GI:5305671"
/protein_id="AAD1767.1"
/translation="MGKRLIVTILALPHRGNAGQCTNFPDIDROSGCCLDID
ECRTIPBACHDMCYNQNCGICLIPFTINVRGPISNPSTISGEPAPAPVPAS
NPFTISPLVCRFGYQMDENQGVDEVDSDHQPNTDICTIEGYSCTDGYW
LLEGCLDIDECRFGYQCOOLCANVGSYSCTCNPFGLNDGSCODVNECTENPCV
QTCVNTGSEFICRDPYELLEDGIIHCSDECSSEFLQHECVNDPQSYGSCPG
VYLDNRSGODINECHRNCTSIQTCVNLGGGFCIDISCEEPYLLIGENRCW
PAHNSCSROPFTILYRDMVYSGRSYPAIDIPQMATRRPGATYITQIKSGNCRF
YKRPSPISATLMTPIKGRPDIDLEMTVNTVINFQSSVIRLITVSYQPF"
BASE COUNT 604 a 656 c 575 g 643 t
ORIGIN
Query Match 82.9%; Score 113.6; DB 10; Length 2478;
Best Local Similarity 89.3%; Pred. No 8.2e-312;
Matches 1200; Conservative 0; Mismatches 144; Indels 0; Gaps 0;
Qy 1 ATCCAGGAATAAAAAGATACCTACTGTTACCATTTGCTGCTCTTCCAAAGCCT 60
Db 320 ATGCCAGATVTAAGAGATACCTACTGTTACCATTTGCTGCTCTTCCAAAGCCT 379
Qy 61 GGGAAATGCAGGACAGTGCAGAGTGGCTTGGACTGAGTGCAGCAGTACAGTGT 120
Db 380 GGGAAATGCAGGACAGTGCAGAGTGGCTTGGACTGAGTGCAGCAGTACAGTGT 439
Qy 121 TTAGATATTGATGAATGCGGAACCATCCCGAGGCGCGGAGAGCATGATGTGT 180
Db 440 CTAGATATTGATGAATGCGGAGACCATCCCTGAGGCTGTGTGGGAGCATGATGTGT 499
Qy 181 AACCAAAATGGGGGTATTATGCAATTCGCCGAGCAAAACCTGTGTATCGAGGGCCTAC 240
Db 500 AACCAAAATGGGGGTATTATGCAATTCGCCGAGCAAAACCTGTGTATCGAGGGCCTAC 559
Qy 241 TCGAACCCCTACTCGACCCCTACTCGAGTGGTGAACCCAGAGTGGCCCAACACTGCA 300
Db 560 TCGAAATCCCTACTCTTACATCTTCTAGAGCCCATCCACAGAGGCGCCCAACACTGCA 619
Qy 301 GCTCAAACTATCCACAGATCTCCAGGCTCTTATATGCGCTTGGATACCAAGATGAT 360
Db 620 GCTTCAACTATCCACAGATCTTCAAGGCGCTCTTGTGTGCGGCTTGGATACCAAGATG 679
Qy 361 GAAACCAACCATGTGTGATGTGAGCAGATGTGCAACAGATTCACCAAGTGCACACCC 420
Db 680 GAAAGCAACCATGTGTGATGTGAGCAGATGTGCAACAGATTCACCAAGTGCACACCC 739
Qy 421 ACCCGATCTGCATCAATCTGAAAGCGGGTACACCTGCTCGACACGAGGATATTTGG 480
Db 740 ACCCGATCTGTATCAACACTGAAAGAGGTTACACCTGCTCGACACGAGGATATTTGG 799
Qy 481 CTTCTGGAAGGCCAGTCTTGAACATTTGATGATGTCGCTATGCTACTGCCACAGCTC 540
Db 800 CTTCTGGAAGGCCAGTCTTGAACATTTGATGATGTCGCTATGCTACTGCCACAGCTC 859
Qy 541 TGTGGAAATGTTCCTGATGCTTATTTGTATCTGACATGCAACCTGTGTTTACCCTCATAG 600
Db 860 TGTGCAAAATGTTCCTGATGCTTATTTGTATCTGACATGCAACCTGTGTTTACCCTCATAG 919
Qy 601 GATGGAAGTCTTGGCAAGATGGAAGGATGTCAGCAGAGAACCCCTGCGCAACAC 660
Db 920 GATGGAAGTCTTGGCAAGATGGAAGGATGTCAGCAGAGAACCCCTGCGTTCACACC 979
Qy 661 TGCCTCAACACCTAGCGCTCTTCAATCTGCGGCTGTGACCCAGAGATGAATCTGAGGAA 720
Db 980 TGTGCAACACCTATGGCTCTTCAATCTGCGGCTGTGACCCAGAGATGAATCTGAGGAA 1039

QY 1021 TGCAGAGACCGCCCTTACCATCTGTACCGGAGATGAGCTGTGTACAGACGCTCC 1080
 DB 1189 TGCAGAGACCGCCCTTACCATCTGTACCGGAGATGAGCTGTGTACAGACGCTCC 1248
 QY 1081 GTTCCCGCTGACATCTTCAAAATGCAAGCAGACCGGCTACCTGTGGGCTATTACAT 1140
 DB 1249 GTTCCCGCTGACATCTTCAAAATGCAAGCAGACCGGCTACCTGTGGGCTATTACAT 1308
 QY 1141 TTCCAGATCAAACTGGGAATGAGGAGAGAAATTTTACATGCGGCAAAAGGCGCCCATC 1200
 DB 1309 TTCCAGATCAAACTGGGAATGAGGAGAGAAATTTTACATGCGGCAAAAGGCGCCCATC 1368
 QY 1201 AGTCCACCTGCTGTGATGACAGACCGCCCATCAAGGCGCCGGAATCCAGCTGACCTG 1260
 DB 1369 AGTCCACCTGCTGTGATGACAGACCGCCCATCAAGGCGCCGGAATCCAGCTGACCTG 1428
 QY 1261 GAAATGATCACTGTCAACACTGTTCATCAACTTCAGAGGAGCTCCGTGATCCGATCGG 1320
 DB 1429 GAAATGATCACTGTCAACACTGTTCATCAACTTCAGAGGAGCTCCGTGATCCGATCGG 1488
 QY 1321 ATATATGTGTGCGAGTACCATTC 1344
 DB 1489 ATATATGTGTGCGAGTACCATTC 1512

RESULT 10
 BC006636 2230 bp mRNA linear ROD 07-ANG-2002
 LOCUS Mus musculus, fibulin 5, clone MGC:5656 IMAGE:3482574, mRNA,
 complete cds.
 ACCESSION BC006636
 VERSION BC006636.1 GI:13879321
 KEYWORDS house mouse.
 SOURCE Mus musculus.
 ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (27-APR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK
 COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgas@remail.nih.gov

Tissue Procurement: Lohar Hennighausen Ph.D., Robin Humphreys
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Hale, S.M.,
 Yoon, V.S., Kovis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.

FEATURES
 source
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 6 Row: f Column: 8
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 6753823.

location/Qualifiers
 1. 2230
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="C57BL/6J"
 /clone="MGC:5656 IMAGE:3482574"
 /tissue="Mammary tumor. WAP-TGF alpha model. 7 months
 old, gross tissue."

CDS
 /clone_11b="NCI_GCAP_Mam5"
 /lab_host="DH10B"
 /note="Vector: pCMV-Sport6"
 60..1406
 /codon_start=1
 /product="Fibulin 5"
 /protein_id="AAH06636.1"
 /db_xref="GI:13879322"
 /db_xref="locusID:23876"

BASE COUNT 562 a 578 c 505 g 585 t
 ORIGIN

Query Match 82.9% Score 1113.6; DB 10; Length 2230;
 Best Local Similarity 89.3%; Pred. No. 8.2e-312;
 Matches 1200; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 1 ATGCCAGGAATAAAGAGTACTACTGTACCATCTGTCTCTCTCCAGCCCT 60
 DB 60 ATGCCAGGATTAAGAGTACTACTGTACCATCTGTCTCTCTCCAGCCCT 119
 QY 61 GGGATGACACAGGACATGACAGCAATGCTTTGACCTGATGCGGACGACAGCTGT 120
 DB 120 GGGATGACACAGGACATGACAGCAATGCTTTGACCTGATGCGGACGACAGCTGT 179
 QY 121 TTGATATTTGATGATGCGGAGCAACATCCCGGCGCTGCGAGAGACATGATGTGT 180
 DB 180 CTGATATTTGATGATGCGGAGCAACATCCCGGCGCTGCGAGAGACATGATGTGT 239
 QY 181 AACCAAAATGCGGGTATTTATGATCCCGGACAAACCTGTGTATGAGGGCCCTAC 240
 DB 240 AACCAAAATGCGGGTATTTATGATCCCGGACAAACCTGTGTATGAGGGCCCTAC 299
 QY 241 TCGAACCCTTCTGACCCCTTACGAGTCCGTACCCAGCAGCTCCCGACACCTCA 300
 DB 300 TCGAACCCTTCTGACCCCTTACGAGTCCGTACCCAGCAGCTCCCGACACCTCA 359
 QY 301 GCTCCAAATCTCCAGCATCTCCAGGCTCTTATATGCGGCTTGTGATACAGATGAT 360
 DB 360 GCTCCAAATCTCCAGCATCTCCAGGCTCTTATATGCGGCTTGTGATACAGATGAT 419
 QY 361 GAAAGCAACCAATGTGTGATGTGAGCAGAGTGTGCAACAGATTCACAGAGTCAACCCC 420
 DB 420 GAAAGCAACCAATGTGTGATGTGAGCAGAGTGTGCAACAGATTCACAGAGTCAACCCC 479
 QY 421 ACCAGATCTGCATCAATCTGAAGGCGGCTACCTGCTCTGACCCAGCATATGTG 480
 DB 480 ACCAGATCTGCATCAATCTGAAGGCGGCTACCTGCTCTGACCCAGCATATGTG 539
 QY 481 CTTTGAAGGCGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 DB 540 CTTTGAAGGCGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 599
 QY 541 TGTGGAATGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 DB 600 TGTGGAATGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 659
 QY 601 GATGGAAGGCTTGTGCAAGATGTGAAGAGTGTGCAAGAGTGTGCAAGAGTGTGCAAGAG 660
 DB 660 GATGGAAGGCTTGTGCAAGATGTGAAGAGTGTGCAAGAGTGTGCAAGAGTGTGCAAGAG 719
 QY 661 TGTGCAACCTTGTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 DB 720 TGTGCAACCTTGTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 779
 QY 721 GATGCGCTTGTGCAAGATGTGAAGAGTGTGCAAGAGTGTGCAAGAGTGTGCAAGAGT 780

QY 901 CTGCACAGACGCTGCTCAATTTACAAAGGGGGCTTCAAAATGATGACGCCATCCGCTGT 960
|||||
Db 1350 CTGCACAGACGCTGCTCAATTTACAAAGGGGGCTTCAAAATGATGACGCCATCCGCTGT 1409
|||||
QY 961 GAGGACCTTATCTGAGGATCAGTATACCGCTGTATGTCTCTGCTGAGAACCTTGGC 1020
|||||
Db 1410 GAGGACCTTATCTGAGGATCAGTATACCGCTGTATGTCTCTGCTGAGAACCTTGGC 1469
|||||
QY 1021 TGCAGAGACCAAGCCCTTACCATCTTGTACCGGGAGCATGAGCTGTGTACAGAGCTTC 1080
|||||
Db 1470 TGCAGAGACCAAGCCCTTACCATCTTGTACCGGGAGCATGAGCTGTGTACAGAGCTTC 1529
|||||
QY 1081 GTTCCGCTGACATCTTCCAAATGCAAGCCAGCCGCTACCCCTGAGGCTTATTCATTT 1140
|||||
Db 1530 GTTCCGCTGACATCTTCCAAATGCAAGCCAGCCGCTACCCCTGAGGCTTATTCATTT 1589
|||||
QY 1141 TTCCAAATCTGAGGATGAGGCGAGAAATTTTACATGCGGCAAAAGGGGCCCTATC 1200
|||||
Db 1590 TTCCAAATCTGAGGATGAGGCGAGAAATTTTACATGCGGCAAAAGGGGCCCTATC 1649
|||||
QY 1201 AGTGCACCTGGTGTATGACACGCCCATCAAAAGGGGCCGGGAAATCCAGCTGAGCTTG 1260
|||||
Db 1650 AGTGCACCTGGTGTATGACACGCCCATCAAAAGGGGCCGGGAAATCCAGCTGAGCTTG 1709
|||||
QY 1261 GAAATGATCAGTGTCAACACTGTTCATCACTTTCAGAGGACCTCCGTGATCCGACTGCGG 1320
|||||
Db 1710 GAAATGATCAGTGTCAACACTGTTCATCACTTTCAGAGGACCTCCGTGATCCGACTGCGG 1769
|||||
QY 1321 ATATATGTGTGCGAGTACCATTC 1344
|||||
Db 1770 ATATATGTGTGCGAGTACCATTC 1793
|||||

RESULT 9
AF093118 2019 bp mRNA 1linear PRI 01-OCT-1998
LOCUS AF093118 Homo sapiens UP50 mRNA, complete cds.
DEFINITION AF093118
VERSION AF093118.1 GI:3676821
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 2019)
AUTHORS Zemel, R., Sholito, O. and Shaul, Y.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-1998) Dept. of Molecular Genetics, Weizmann
Institute of Science, Rehovot 76100, Israel
FEATURES
source location/Qualifiers
1..2019
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="urine"
1..168
159..1515
/note="50KD: similar to human SI-5 encoded by GenBank
Accession Number U03877; possible membrane and
extracellular matrix protein; urine p50 protein"
/codon_start=1
/product="UP50"
/protein_id="AAC62107.1"
/db_xref="GI:3676822"
/translation="MGIRILITVTLALILPSGNAOACTNGFDLDROSGCIDI
ECRTPEACRBDAMCVNONGYILCHSTNPISPISTPGIPAPAPLISAP
NPTISRLICRFQYMDENQCVDECAIDSHQNPQICIMKGGYTSCIDGY
LLEGQCLIDECRYGCOQLCANVGSYSCINPFTLNEBGRSCODNECATENPC
QTCVNTYGSFICRDEGYLEEDGVHSCSDMECSFELCOHECVNOGTYECSPPG
YILDNRSQODINECEHRHNTCNLOTCYNLOGFKCIDIRCEPEPLRIISDNRCM
PAENPCRDOPFTILYRDMVSVGRSYPADIFOMQATRYGAYIYIIOIKSNBER
YKROTPISATILMTMRPIKGPRIQLDLEMITVNTVINFKSSVIRLITVSYF"

3'UTR 3'UTR
BASE COUNT 478 a 585 c 490 g 465 t 1 others

ORIGIN
Query Match 98.8% Score 1328: DB 9: Length 2019;
Best Local Similarity 99.3%: Pred. No. 0; Mismatches 10; Indels 0; Gaps 0;
Matches 1334; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 ATCCAGAGATATAAAGATATCTACCTGTACATTTCTGCGCTCTGTTCCAAAGCCT 60
Db 169 ATCCAGAGATATAAAGATATCTACCTGTACATTTCTGCGCTCTGTTCCAAAGCCT 228
QY 61 GGGATATGACAGGACAGTGCAGAGATGCTTTGACCTGTGATGCCAGTACAGAGTGT 120
Db 229 GGGATATGACAGGACAGTGCAGAGATGCTTTGACCTGTGATGCCAGTACAGAGTGT 288
QY 121 TTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
Db 289 TTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 348
QY 181 AACCAAAATGCGGGGATTTTATGATTCGCCGACAAACCTGTATGAGGGCCCTAC 240
Db 349 AACCAAAATGCGGGGATTTTATGATTCGCCGACAAACCTGTATGAGGGCCCTAC 408
QY 241 TCGAACCCTACTGACCCCTACTCAGTCCGTACCCAGACAGCTGCCACACTGTCA 300
Db 409 TCGAACCCTACTGACCCCTACTCAGTCCGTACCCAGACAGCTGCCACACTGTCA 468
QY 301 GCTCCAAATCTCCAGATCTCCAGGCTCTTATGATGCGCTTGGATACAGATGAT 360
Db 469 GCTCCAAATCTCCAGATCTCCAGGCTCTTATGATGCGCTTGGATACAGATGAT 528
QY 361 GAAAGCAACAAATGTGTGATGTGAGAGAGTGTGCACAGATTTCCACAGTGCACACC 420
Db 529 GAAAGCAACAAATGTGTGATGTGAGAGAGTGTGCACAGATTTCCACAGTGCACACC 588
QY 421 ACCCAGATCTGCATCAATPACTGAAGCGGGTACACCTGCTCTCGACAGAGGATTTGG 480
Db 589 ACCCAGATCTGCATCAATPACTGAAGCGGGTACACCTGCTCTCGACAGAGGATTTGG 648
QY 481 CTTCTGGAAGGCGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Db 649 CTTCTGGAAGGCGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 708
QY 541 TGTGCAATGTCTCTGATCTTATCTTGTACATGCAACCTGTGTTTACCCTCATGAG 600
Db 709 TGTGCAATGTCTCTGATCTTATCTTGTACATGCAACCTGTGTTTACCCTCATGAG 768
QY 601 GATGGAAGTCTTGTCCAAAGATGTGAACAGTGTGCGACCGAGAACCCCTGCGTCAAC 660
Db 769 GATGGAAGTCTTGTCCAAAGATGTGAACAGTGTGCGACCGAGAACCCCTGCGTCAAC 828
QY 661 TGCGTCAACACCTACGCTCTTTCATCTGCGCGGTGTGACCGAGATATGAATTGAGGA 720
Db 829 TGCGTCAACACCTACGCTCTTTCATCTGCGCGGTGTGACCGAGATATGAATTGAGGA 888
QY 721 GATGGGCTTATGCTACATGATGAGAGAGTGCAGCTTCTGAGTTCTCTGCGCAACAT 780
Db 889 GATGGGCTTATGCTACATGATGAGAGAGTGCAGCTTCTGAGTTCTCTGCGCAACAT 948
QY 781 GAGTGTGTGAACAGCCCGGACATCTTGTCTCTGCGCTCCAGAGCTTACCTCTGCTG 840
Db 949 GAGTGTGTGAACAGCCCGGACATCTTGTCTCTGCGCTCCAGAGCTTACCTCTGCTG 1008
QY 841 GATGACAACCGAAGCTCCCAAGCATCAAGCAAGTGTGACAGGAACACACAGTGCAC 900
Db 1009 GATGACAACCGAAGCTCCCAAGCATCAAGCAAGTGTGACAGGAACACACAGTGCAC 1068
QY 901 CTGCACAGACGCTGCTCAATTTTCAAGGGGGCTTCAAAATGATGACGCCATCCGCTGT 960
Db 1069 CTGCACAGACGCTGCTCAATTTTCAAGGGGGCTTCAAAATGATGACGCCATCCGCTGT 1128
QY 961 GAGGACCTTATCTGAGGATCAGTATACCGCTGTATGTCTCTGCTGAGAACCTTGGC 1020
Db 1129 GAGGACCTTATCTGAGGATCAGTATACCGCTGTATGTCTCTGCTGAGAACCTTGGC 1188

QY 601 GATGGAAGGCTCTGCCAAGATGTGAACAGAGTGTCCACCGAAGAACCCCTGGCTGCAACC 660
 DB 1050 GATGGAAGGCTCTGCCAAGATGTGAACAGAGTGTCCACCGAAGAACCCCTGGCTGCAACC 1109
 QY 661 TGCCTGCAACACCTACGAGCTCTTTCATCTGCCGCTGTGACCCAGATATGAATGAGGAA 720
 DB 1110 TGCCTGCAACACCTACGAGCTCTTTCATCTGCCGCTGTGACCCAGATATGAATGAGGAA 1169
 QY 721 GATGGGCTTCATTGCAAGTGTATGAGAGTGCAGCTTCTGTGACTTCTGTCCACAT 780
 DB 1170 GATGGGCTTCATTGCAAGTGTATGAGAGTGCAGCTTCTGTGACTTCTGTCCACAT 1229
 QY 781 GAGTGTGTAACACGAGCCGCGACATACCTTGTGCTCCGCTCCAGGCTACATCTGCG 840
 DB 1230 GAGTGTGTAACACGAGCCGCGACATACCTTGTGCTCCGCTCCAGGCTACATCTGCG 1289
 QY 841 GATGACACCGAAGCTGTCCAAAGACATCAAGAAATGTAGAGCAGAGAACACAGCTGCAAC 900
 DB 1290 GATGACACCGAAGCTGTCCAAAGACATCAAGAAATGTAGAGCAGAGAACACAGCTGCAAC 1349
 QY 901 CTGCAGCAACACGCTGTACATTTTCAAGGGGGCTTCAAAATGATGACCCCATCCGCTGT 960
 DB 1350 CTGCAGCAACACGCTGTACATTTTCAAGGGGGCTTCAAAATGATGACCCCATCCGCTGT 1409
 QY 961 GAGAGACCTTATCTGAGATCACTGATTAACCGCTGTATGTCTGCTGAGAACCCCTGAC 1020
 DB 1410 GAGAGACCTTATCTGAGATCACTGATTAACCGCTGTATGTCTGCTGAGAACCCCTGAC 1469
 QY 1021 TGCAGAGACACGAGCCCTTTTACCATCTGTGTACCGGGACATGACGCTGTGTGAGAGCTTC 1080
 DB 1470 TGCAGAGACACGAGCCCTTTTACCATCTGTGTGTACCGGGACATGACGCTGTGTGAGAGCTTC 1529
 QY 1081 GTTCCCGCTGACATCTTCCAAATGCAAGCAGCAGCCGCTACCGTGGGGCTTATACAT 1140
 DB 1530 GTTCCCGCTGACATCTTCCAAATGCAAGCAGCAGCCGCTACCGTGGGGCTTATACAT 1589
 QY 1141 TTCCAGATCAAAATCTGGGAATGAGGGCAGAGAAATTTTACATGCGGCAAAAGGGCCCATC 1200
 DB 1590 TTCCAGATCAAAATCTGGGAATGAGGGCAGAGAAATTTTACATGCGGCAAAAGGGCCCATC 1649
 QY 1201 AGTGCACACCTGCTGTATGACAGCCGCTCATAAAGGGCCCGGAAATCAGCTGACCTTG 1260
 DB 1650 AGTGCACACCTGCTGTGTATGACAGCCGCTCATAAAGGGCCCGGAAATCAGCTGACCTTG 1709
 QY 1261 GAAATGATCACTGTCAACACTGTATCATCACTTGTGAGGAGCAGCTCGGTATCCGACTGG 1320
 DB 1710 GAAATGATCACTGTCAACACTGTATCATCACTTGTGAGGAGCAGCTCGGTATCCGACTGG 1769
 QY 1321 ATATATGTGTCCAGTACCCATTC 1344
 DB 1770 ATATATGTGTCCAGTACCCATTC 1793
 RESULT 8
 AX403659 2609 bp DNA linear PAT 14-JUN-2002
 LOCUS AX403659
 DEFINITION Sequence 14 from Patent WO0077037.
 ACCESSION AX403659
 VERSION AX403659.1 GI:21437116
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS 1 Ashkenazi, A., Baker, K., Botstein, D., Desnoyers, L., Eaton, D.L.,
 Ferrara, N., Fong, S., Gao, W.Q., Gerber, H., Gertlisen, M.E.,
 Godard, A., Godowski, P., Gurney, A., Kijavlin, I.J., Mather, J.,
 Napier, M., Pan, J., Paoni, N., Roy, M., Tamas, D., Watanabe, C.,
 Williams, P.M., Wood, W.I. and Zhang, Z.
 TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
 the same

JOURNAL Patent: WO 0077037-A 14 21-DEC-2000;
 Genentech Inc. (US)
 FEATURES Location/Qualifiers
 source 1..2609 /organism="Homo sapiens"
 BASE COUNT 622 a 718 c 622 g 647 t
 ORIGIN
 Query Match 99.98; Score 1342.4; DB 6; Length 2609;
 Best Local Similarly 99.98; Pred. No. 0;
 Matches 1343; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGCCAGGAATTAAGAAAGATACATCAGTGTACATTCCTGCTGTCTGCTTCCAGCCCT 60
 DB 450 ATGCCAGGAATTAAGAAAGATACATCAGTGTACATTCCTGCTGTCTGCTTCCAGCCCT 509
 QY 61 GGAATGACACAGGACAGTGCAGCAATGCTTTGACCTGTGATCCGCACTCAGACAGTGT 120
 DB 510 GGAATGACACAGGACAGTGCAGCAATGCTTTGACCTGTGATCCGCACTCAGACAGTGT 569
 QY 121 TTGATATTTATGATATGATCCGCAACCATCCCGAGGCTTCCGAGAGACATATGTGT 180
 DB 570 TTGATATTTATGATATGATCCGCAACCATCCCGAGGCTTCCGAGAGACATATGTGT 629
 QY 181 AACCAAAATGGGGATTTTATGATTCGCCGACAAACCTGTGTATGAGAGGCCCTTAC 240
 DB 630 AACCAAAATGGGGATTTTATGATTCGCCGACAAACCTGTGTATGAGAGGCCCTTAC 689
 QY 241 TCGAACCCTTACTGTGACCCCTTACTCAGTCCGTATCCAGACAGCTGCCCCACACTGTCA 300
 DB 690 TCGAACCCTTACTGTGACCCCTTACTCAGTCCGTATCCAGACAGCTGCCCCACACTGTCA 749
 QY 301 GCTCAAAATATCCAGATATCCAGATCTCCAGGCTCTTATATGACCGCTTGTGATACAGATGAT 360
 DB 750 GCTCAAAATATCCAGATATCCAGATCTCCAGGCTCTTATATGACCGCTTGTGATACAGATGAT 809
 QY 361 GAAAGCAACCAATGTGTGATGTGAGAGTGTGCAACAGATTTCCACAGTGCACACCC 420
 DB 810 GAAAGCAACCAATGTGTGATGTGAGAGTGTGCAACAGATTTCCACAGTGCACACCC 869
 QY 421 ACCCAATGTGTGATCTGAAAGGGGGGTATACACTCTCTCTGACAGGAGATTTGG 480
 DB 870 ACCCAATGTGTGATCTGAAAGGGGGGTATACACTCTCTCTGACAGGAGATTTGG 929
 QY 481 CTTCTGGAAGGCAAGCTGTAGACATGATTAATGTGCTATGTTACTGCGAGAGCTC 540
 DB 930 CTTCTGGAAGGCAAGCTGTAGACATGATTAATGTGCTATGTTACTGCGAGAGCTC 989
 QY 541 TGTGCAATGTCTCTGATCTTATCTTGTATGATCAACACCTGTGTTTACCTCAATGAG 600
 DB 990 TGTGCAATGTCTCTGATCTTATCTTGTATGATCAACACCTGTGTTTACCTCAATGAG 1049
 QY 601 GATGGAAGGCTCTGCCAAGATGTGAACAGAGTGTCCACCGAAGAACCCCTGGCTGCAACC 660
 DB 1050 GATGGAAGGCTCTGCCAAGATGTGAACAGAGTGTCCACCGAAGAACCCCTGGCTGCAACC 1109
 QY 661 TGCCTGCAACACCTACGAGCTCTTTCATCTGCCGCTGTGACCCAGATATGAATGAGGAA 720
 DB 1110 TGCCTGCAACACCTACGAGCTCTTTCATCTGCCGCTGTGACCCAGATATGAATGAGGAA 1169
 QY 721 GATGGGCTTCATTGCAAGTGTATGAGAGTGCAGCTTCTGTGACTTCTGTCCACAT 780
 DB 1170 GATGGGCTTCATTGCAAGTGTATGAGAGTGCAGCTTCTGTGACTTCTGTCCACAT 1229
 QY 781 GAGTGTGTAACACGAGCCGCGACATACCTTGTGCTCCGCTCCAGGCTACATCTGCG 840
 DB 1230 GAGTGTGTAACACGAGCCGCGACATACCTTGTGCTCCGCTCCAGGCTACATCTGCG 1289
 QY 841 GATGACACCGAAGCTGTCCAAAGACATCAAGAAATGTAGAGCAGAGAACACAGCTGCAAC 900
 DB 1290 GATGACACCGAAGCTGTCCAAAGACATCAAGAAATGTAGAGCAGAGAACACAGCTGCAAC 1349

OY	241	TCGACCCCTACTGACACCCCTACTGAGTCCCTACCCAGACACTGCCACCACCTCTCA	300
Db	622	TCGACCCCTACTGACACCCCTACTGAGTCCCTACTACCCAGACACTGCCACCACCTCTCA	681
OY	301	GCTCCAAATATCCACGATCTCTCAGGCGCTCTATATGCGCGCTTTGGATACCGATGGAT	360
Db	682	GCTCCAAATATCCACGATCTCTCAGGCGCTCTATATGCGCGCTTTGGATACCGATGGAT	741
OY	361	GAAGAACAACCAATGTGTGGATGTGGACAGGTGTGCACAGATATCCACAGTGCAACCC	420
Db	742	GAAGAACAACCAATGTGTGGATGTGGACAGGTGTGCACAGATATCCACAGTGCAACCC	801
OY	421	ACCCAGATCTGCATCAATACTGAAAGCGGGTACACCTGCTCCGACCGACGAGATATG	480
Db	802	ACCCAGATCTGCATCAATACTGAAAGCGGGTACACCTGCTCCGACCGACGAGATATG	861
OY	481	CTTTCTGGAAGGCCAGTCTTTAGACATATGATGATGTGCTATGATTACTGCCAGAGCTC	540
Db	862	CTTTCTGGAAGGCCAGTCTTTAGACATATGATGATGTGCTATGATTACTGCCAGAGCTC	921
OY	541	TGAGCAATGTTCCCTGGATCCTATCTTGTACATGACACCGTGTTTACCTCAATGAG	600
Db	922	TGAGCAATGTTCCCTGGATCCTATCTTGTACATGACACCGTGTTTACCTCAATGAG	981
OY	601	GATGGAAGTCTTGCCAAAGATGTGAACAGTGTGCCACAGAGAACCCCTGCTGCAAAAC	660
Db	982	GATGGAAGTCTTGCCAAAGATGTGAACAGTGTGCCACAGAGAACCCCTGCTGCAAAAC	1041
OY	661	TGCGTCANACCTACGAGCTCTTTCATCTTCCCGCTGTGACCCAGATATGAACTTAGAA	720
Db	1042	TGCGTCANACCTACGAGCTCTTTCATCTTCCCGCTGTGACCCAGATATGAACTTAGAA	1101
OY	721	GATGCGCTTCATTCGACAGTATGTGACAGATGCACCTCTCTGATCTCTCGCCACAT	780
Db	1102	GATGCGCTTCATTCGACAGTATGTGACAGATGCACCTCTCTGATCTCTCGCCACAT	1161
OY	781	GAGTGTGTGAACACACCCGGACATATCTTGCTCTGCTCCCTCCAGGCTACATCTGCTG	840
Db	1162	GAGTGTGTGAACACACCCGGACATATCTTGCTCTGCTCCCTCCAGGCTACATCTGCTG	1221
OY	841	GATGCAACCCGAAGCTGGCCAAAGACATCAACGAATGTGACAGACAGAACACACAGTGCAC	900
Db	1222	GATGCAACCCGAAGCTGGCCAAAGACATCAACGAATGTGACAGACAGAACACACAGTGCAC	1281
OY	901	CTGCAGCAGAGCTGCTACAAATTTACAGAGGGGCTTCAATGCATGCAGCCCATCGCTGT	960
Db	1282	CTGCAGCAGAGCTGCTACAAATTTACAGAGGGGCTTCAATGCATGCAGCCCATCGCTGT	1341
OY	961	GAGGAGCCTTATCTGAGGATCAGTATATACCGCTGATGTCTCTGCTGAGAACCCCTGGC	1020
Db	1342	GAGGAGCCTTATCTGAGGATCAGTATATACCGCTGATGTCTCTGCTGAGAACCCCTGGC	1401
OY	1021	TGCAAGACACAGCCCTTTACCATCTTGTATCCGGGACATGACAGTGTGTCAGAGCTTCC	1080
Db	1402	TGCAAGACACAGCCCTTTACCATCTTGTATCCGGGACATGACAGTGTGTCAGAGCTTCC	1461
OY	1081	GTTCCCGGTGATCTTCCAAATGCAAGCAGCAGCCCGCTACCTGGGGCTATTACATT	1140
Db	1462	GTTCCCGGTGATCTTCCAAATGCAAGCAGCAGCCCGCTACCTGGGGCTATTACATT	1521
OY	1141	TTCCAGATCAAAATCTGGGAATGAGGGCAGAAATTTACATGCGGCAAAAGGGGCCCATC	1200
Db	1522	TTCCAGATCAAAATCTGGGAATGAGGGCAGAAATTTACATGCGGCAAAAGGGGCCCATC	1581
OY	1201	AGTCCACCCCTGGTGTATGACACGCCCCCATCAAGAGGGCCCGGAAATCCAGCTGACTTG	1260
Db	1582	AGTCCACCCCTGGTGTATGACACGCCCCCATCAAGAGGGCCCGGAAATCCAGCTGACTTG	1641
OY	1261	GAATGATCACTGTGCACACGTGCATCAACTGAGAGGCAAGCTCCGTGATCCGACATGCGG	1320
Db	1642	GAATGATCACTGTGCACACGTGCATCAACTGAGAGGCAAGCTCCGTGATCCGACATGCGG	1701
OY	1321	ATAATGTGTGCGAGTACCCATTG	1344

Db	1702	ATATATGTTGTCGACGATCCATTC	1725
RESULT 7			
LOCUS	AX376340	2609 bp	DNA
DEFINITION	Sequence 407 from Patent WO0168848.		linear
ACCESSION	AX376340		PAT 01-MAR-2002
VERSION	AX376340.1	GI:19170561	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Baker, K.P., Chen, J., Desnoyers, L., Goddard, A., Godowski, P.J., Zhang, Z., Guney, A.L., Pan, J., Smith, V., Matanabe, C.K., Wood, W.T. and		
TITLE	Secreted and transmembrane polypeptides and nucleic acids encoding the same		
JOURNAL	Patent: WO 0168848-A 407 20-SEP-2001;		
FEATURES	Genentech, Inc. (US)		
SOURCE	location/qualifiers		
	1..2609		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	622 a 718 c 622 g 647 t		
ORIGIN			
Query Match	99.98%	Score 1342.4;	DB 6; Length 2609;
Best Local Similarity	99.98%	Pred. No. 0;	
Matches 1343;	Conservative 0;	Mismatches 1;	Indels 0; Gaps 0;
QY	1	ATGCCAGAAATAAAGATACATACATGTTACATTCCTGGCTGCTGCTTCCAAAGCCCT	60
Db	450	ATGCCAGAAATAAAGATACATACATGTTACATTCCTGGCTGCTGCTTCCAAAGCCCT	509
QY	61	GGGAATGCACAGGCACAGTGCACGATGCTTTGACCTGGATGCCAGTCAGCAGATGT	120
Db	510	GGGAATGCACAGGCACAGTGCACGATGCTTTGACCTGGATGCCAGTCAGCAGATGT	569
QY	121	TTAGATATTTGATGATATCCGACACATCCCGAGGCCGCGGAGAGACATGATGTGTT	180
Db	570	TTAGATATTTGATGATATCCGACACATCCCGAGGCCGCGGAGAGACATGATGTGTT	629
QY	181	AACCAAAATGGCGGGATTTTATGATATCCCGGACAAACCTGTATGAGAGGCCCTAC	240
Db	630	AACCAAAATGGCGGGATTTTATGATATCCCGGACAAACCTGTATGAGAGGCCCTAC	689
QY	241	TGCAACCCCTACTGCACCCCTCTACTGATGCGGTACCCAGACACTGCCACACCTCA	300
Db	690	TGCAACCCCTACTGCACCCCTCTACTGATGCGGTACCCAGACACTGCCACACCTCA	749
QY	301	GCTCCAAATCTATCCACGATCTCCAGGCCCTTATATGCGCTTTGGATACGATGAT	360
Db	750	GCTCCAAATCTATCCACGATCTCCAGGCCCTTATATGCGCTTTGGATACGATGAT	809
QY	361	GAAAGAACCAATGTGTGATGTGGACAGAGTGTGGCAACAGATTCCACAGTGCACACCC	420
Db	810	GAAAGAACCAATGTGTGATGTGGACAGAGTGTGGCAACAGATTCCACAGTGCACACCC	869
QY	421	ACCAGATCTGCATCAATACTGTAAGCGGGGTACACTGTCTCCGACGACGAGATTTGG	480
Db	870	ACCAGATCTGCATCAATACTGTAAGCGGGGTACACTGTCTCCGACGACGAGATTTGG	929
QY	481	CTTCTGGAAGGCCAGTCTTAGACATGATGAATGTGCTATGTTACTGCAGACGCTC	540
Db	930	CTTCTGGAAGGCCAGTCTTAGACATGATGAATGTGCTATGTTACTGCAGACGCTC	989
QY	541	TGTGCAATGTTCTGATGCTATTTCTTATATGCAACCGGGTTTAACTCCCTCAATAG	600
Db	990	TGTGCAATGTTCTGATGCTATTTCTTATATGCAACCGGGTTTAACTCCCTCAATAG	1049

AUTHORS Bandman, O., Corley, N.C. and Guegler, K.J.
 TITLE Human extracellular matrix proteins
 JOURNAL Patent: US 5872234-A 2 16-FEB-1999;
 FEATURES Location/Qualifiers
 Source 1..2550
 /organism="unknown"
 BASE COUNT 623 a 690 c 596 g 641 t
 ORIGIN

Query Match 100.0%; Score 1344; DB 6; Length 2550;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 ATGCCAGGATTAAGAGATCTACGTATACATTCGCTCTCTGCTTCCAGCCCT 60
DB 382 ATGCCAGGATTAAGAGATCTACGTATACATTCGCTCTCTGCTTCCAGCCCT 441
OY 61 GGGAAATGCACAGGACAGTGCACGAAATGGCTTTCAGCTGCAGTCCAGTACAGTGT 120
DB 442 GGGAAATGCACAGGACAGTGCACGAAATGGCTTTCAGCTGCAGTCCAGTACAGTGT 501
OY 121 TTAGATATTGATGAATATGCCAATCCCGAGGCTTCGAGAGACATGATGTGT 180
DB 502 TTAGATATTGATGAATATGCCAATCCCGAGGCTTCGAGAGACATGATGTGT 561
OY 181 AACCAAAATGGCGGATTTATGATCTCCCGGACAAACCTGTGTATCGAGGCGCTAC 240
DB 562 AACCAAAATGGCGGATTTATGATCTCCCGGACAAACCTGTGTATCGAGGCGCTAC 621
OY 241 TCGAACCCCTACTGACACCCCTACTACAGTTCGCTACCCAGAGCTGCCACACCTCTCA 300
DB 622 TCGAACCCCTACTGACACCCCTACTACAGTTCGCTACCCAGAGCTGCCACACCTCTCA 681
OY 301 GCTCCAACTATCCACGATCTCCAGGCTCTTATATGCGCTTGGATACAGATGAT 360
DB 682 GCTCCAACTATCCACGATCTCCAGGCTCTTATATGCGCTTGGATACAGATGAT 741
OY 361 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAACAGATTTCCACAGTGCNAACCC 420
DB 742 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAACAGATTTCCACAGTGCNAACCC 801
OY 421 ACCCAGATCTGCATCAATCTGAAAGCGGGTACACCTGCTCTCGACCCAGCGATTTGG 480
DB 802 ACCCAGATCTGCATCAATCTGAAAGCGGGTACACCTGCTCTCGACCCAGCGATTTGG 861
OY 481 CTCTTGAAGGCGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
DB 862 CTCTTGAAGGCGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 921
OY 541 TGTGCGAATGTCTGAGATCTATCTTGTATGATGCAACCTGTGTTTACCTCAATGAG 600
DB 922 TGTGCGAATGTCTGAGATCTATCTTGTATGATGCAACCTGTGTTTACCTCAATGAG 981
OY 601 GATGGAAGTCTTTCGCAAGATGTGAGAGAGTGTGCGACCGGAACCCCTGGGTGCAAC 660
DB 982 GATGGAAGTCTTTCGCAAGATGTGAGAGAGTGTGCGACCGGAACCCCTGGGTGCAAC 1041
OY 661 TCGCTCAACACCTACGCGCTTTCATCTGCGCTGTACACCGAGATATGAACTTGAGAA 720
DB 1042 TCGCTCAACACCTACGCGCTTTCATCTGCGCTGTACACCGAGATATGAACTTGAGAA 1101
OY 721 GATGCGCTTCATTCAGTGTATGAGAGAGTGTGCGCTTCTGTGAGTTCCTGTGCCAAT 780
DB 1102 GATGCGCTTCATTCAGTGTATGAGAGAGTGTGCGCTTCTGTGAGTTCCTGTGCCAAT 1161
OY 781 GATGTGTATACACGAGCGCGGACATATCTGTCTCCGCGCTCCAGGATATCTGTGTG 840
DB 1162 GATGTGTATACACGAGCGCGGACATATCTGTCTCCGCGCTCCAGGATATCTGTGTG 1221
OY 841 GATGACAAACGAGTGTGCAAGACATCAAGATGTGAGACAGAAACCAACAGTGCAC 900
DB 1222 GATGACAAACGAGTGTGCAAGACATCAAGATGTGAGACAGAAACCAACAGTGCAC 1281

```

```

OY 901 CTGACAGCAGTGTCTCAATTTTACAGGGGCTTCAATGATGACGCCATCCGCTGT 960
DB 1282 CTGACAGCAGTGTCTCAATTTTACAGGGGCTTCAATGATGACGCCATCCGCTGT 1341
OY 961 GAGAGGCTTATCTGAGGATCAGTAAACCGCTGTATGTGTCTGTGAGAACCTGTGC 1020
DB 1342 GAGAGGCTTATCTGAGGATCAGTAAACCGCTGTATGTGTCTGTGAGAACCTGTGC 1401
OY 1021 TGCAGAGACACGCCCTTACCATTCTTGTACCGGACATGAGAGTGTGTGACAGGCTGC 1080
DB 1402 TGCAGAGACACGCCCTTACCATTCTTGTACCGGACATGAGAGTGTGTGACAGGCTGC 1461
OY 1081 GTTCCGCTGACATCTTCCAAATGCAAGCCACGACCCGCTACCCCTGGGGCTATTACAT 1140
DB 1462 GTTCCGCTGACATCTTCCAAATGCAAGCCACGACCCGCTACCCCTGGGGCTATTACAT 1521
OY 1141 TTCCAGATCAATCTGGAATGAGGCAAGAAATTTTACATGCGGCAACGGGCCCATC 1200
DB 1522 TTCCAGATCAATCTGGAATGAGGCAAGAAATTTTACATGCGGCAACGGGCCCATC 1581
OY 1201 AGTCCACCCCTGTGATGACACGCGCCCATCAAGGCGCCCGGAAATCCAGCTGACCTG 1260
DB 1582 AGTCCACCCCTGTGATGACACGCGCCCATCAAGGCGCCCGGAAATCCAGCTGACCTG 1641
OY 1261 GAAATGATCACTGTCAACACTGTATCAACTTCAAGAGGACAGTCCGTGATCCGACTGCG 1320
DB 1642 GAAATGATCACTGTCAACACTGTATCAACTTCAAGAGGACAGTCCGTGATCCGACTGCG 1701
OY 1321 ATATATGTGTGCGAGTACCATTC 1344
DB 1702 ATATATGTGTGCGAGTACCATTC 1725

```

RESULT 6
 AR173204
 LOCUS AR173204 2550 bp DNA linear PAT 17-DEC-2001
 DEFINITION Sequence 2 from patent US 6303765.
 ACCESSION AR173204
 VERSION AR173204.1 GI:17912695
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 2550)
 AUTHORS Bandman, O., Corley, N.C. and Guegler, K.J.
 TITLE Human extracellular matrix proteins
 JOURNAL Patent: US 6303765-A 2 16-OCT-2001;
 FEATURES Location/Qualifiers
 Source 1..2550
 /organism="unknown"
 BASE COUNT 623 a 690 c 596 g 641 t
 ORIGIN

Query Match 100.0%; Score 1344; DB 6; Length 2550;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 ATGCCAGGATTAAGAGATCTACGTATACATTCGCTCTCTGCTTCCAGCCCT 60
DB 382 ATGCCAGGATTAAGAGATCTACGTATACATTCGCTCTCTGCTTCCAGCCCT 441
OY 61 GGGAAATGCACAGGACAGTGCACGAAATGGCTTTCAGCTGCAGTCCAGTACAGTGT 120
DB 442 GGGAAATGCACAGGACAGTGCACGAAATGGCTTTCAGCTGCAGTCCAGTACAGTGT 501
OY 121 TTAGATATTGATGAATATGCCAATCCCGAGGCTTCGAGAGACATGATGTGT 180
DB 502 TTAGATATTGATGAATATGCCAATCCCGAGGCTTCGAGAGACATGATGTGT 561
OY 181 AACCAAAATGGCGGATTTATGATCTCCCGGACAAACCTGTGTATCGAGGCGCTAC 240
DB 562 AACCAAAATGGCGGATTTATGATCTCCCGGACAAACCTGTGTATCGAGGCGCTAC 621

```


source
1. .2328
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="14"
map:14q32.1"
159. .1515
/note="DANCE: contains 6 cBEGF domains and 1 RGD motif;
secreted protein; expressed in developmental aorta, neural
crest cells, balloon injured vessels, atherosclerotic
lesions"
/codon_start=1
/product="developmental arteries and neural crest EGF-like
protein"
/protein_id="AAD1768.1"
/db_xref="GI:5305673"
/translation="MGKIRLITVTILALPSRGNNOACTNFDLDROSGCLDID
ERTIPACRDMCYNONGYICLIPINFTYKSPNPSYSPISGPIPAAPLSAP
NPPTISRLICRGYQMDENOCVDDECATDHSQNPOTICINTEGPGSCIDGW
LLEGOLDIDECRYGCOOLCANVPSGSCNPFILNEDGSCODVNECATENPCV
QTCVNTYGSFICDDEYELDEGVSCHSDDECSFSELCQHECVNODYFSCSPG
YLLDNRSCODINECEHRNHTCNLOOTCYNLOGKFCIDIRCEPEYLRISDNRCM
PENPCRDOPFTILYRDMDVYGRSVPADIFOMATRYRGAYYIPIQIKSGNEREF
YMRQNPISATLMTVRPIKPREIQLDLEMTVTNTVINFKSSVIRLRITYSQIF"

BASE COUNT 560 a 646 c 541 g 581 t
ORIGIN

Query Match 100.0%; Score 1344; DB 9; Length 2328;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGCCAGAAATATAAGATATCTACTGTATACATCTGCTCTGTCTTCCAAGCCCT 60
169 ATGCCAGAAATATAAGATATCTACTGTATACATCTGCTCTGTCTTCCAAGCCCT 228
61 GGGAAATGACAGGACAGTGCAGAAATGCTTACCTGATCCGCACTGAGACAGTGT 120
229 GGGAAATGACAGGACAGTGCAGAAATGCTTACCTGATCCGCACTGAGACAGTGT 288
121 TTAGATATGATGAAATGCGCAACATCCCGAGGCTGCGCAGAGAGACATATGTGT 180
289 TTAGATATGATGAAATGCGCAACATCCCGAGGCTGCGCAGAGAGACATATGTGT 348
181 AAACCAATATGGCGGGTATTTATGCAATCCCGGACAAACCTGTGTATGAGGGCCCTAC 240
349 AAACCAATATGGCGGGTATTTATGCAATCCCGGACAAACCTGTGTATGAGGGCCCTAC 408
241 TGGAAACCCCTACTGCAACCCCTACTGAGTCCGTAACCCAGAGAGTCCGCACTGCA 300
409 TGGAAACCCCTACTGCAACCCCTACTGAGTCCGTAACCCAGAGAGTCCGCACTGCA 468
301 GGTCCAAATATCCCAAGATCTCAGAGCTCTTATATGCGCTTGGATACAGATGGAT 360
469 GGTCCAAATATCCCAAGATCTCAGAGCTCTTATATGCGCTTGGATACAGATGGAT 528
361 GAAAGCAACCAATGTGTGATGTGAGAGAGTGCACACAGATTTCCACACAGTCAACCC 420
529 GAAAGCAACCAATGTGTGATGTGAGAGAGTGCACACAGATTTCCACACAGTCAACCC 588
421 ACCCAAGATGTGATCAATATGAAAGGGGGGTACACCTGCTCTGACGACGAGGATTTGG 480
589 ACCCAAGATGTGATCAATATGAAAGGGGGGTACACCTGCTCTGACGAGGATTTGG 648
481 CTCTGGAAGGCGAGTGTAGACATGATGATGATGATGATGATGATGATGATGATGAT 540
649 CTCTGGAAGGCGAGTGTAGACATGATGATGATGATGATGATGATGATGATGATGAT 708
541 TGTGGAATGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
709 TGTGGAATGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 768
601 GATGGAAGGTCTTGGCAAGATGTGAGAGAGTGTGCAACCCGAGAACCCCTGTGCAAAAC 660
769 GATGGAAGGTCTTGGCAAGATGTGAGAGAGTGTGCAACCCGAGAACCCCTGTGCAAAAC 828

661 TGGCTCAACACCTTACGCGCTTTTCATCTGCGCTGTGACCCAGATATGAATTGAGAA 720
829 TGGCTCAACACCTTACGCGCTTTTCATCTGCGCTGTGACCCAGATATGAATTGAGAA 888
721 GATGCGGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
889 GATGCGGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 948
781 GATGCGGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
949 GATGCGGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1008
841 GATGCGGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
1009 GATGCGGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1068
901 CTGACAGACAGCGTGTCAATTTAACAAGGGGGCTTCAATGCAATGCAATGCAATGCA 960
1069 CTGACAGACAGCGTGTCAATTTAACAAGGGGGCTTCAATGCAATGCAATGCAATGCA 1128
961 GATGCGGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
1129 GATGCGGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1188
1021 TGGAGAGACAGCGCTTACCATCTTGTACCGGGGACATGAGAGTGTGACAGAGCTTC 1080
1189 TGGAGAGACAGCGCTTACCATCTTGTACCGGGGACATGAGAGTGTGACAGAGCTTC 1248
1081 GTTCCGCGTACATCTTCCAAATGCAAGCCAGACCGCTTACCGTGGGGCTTATTAAT 1140
1249 GTTCCGCGTACATCTTCCAAATGCAAGCCAGACCGCTTACCGTGGGGCTTATTAAT 1308
1141 TTTCCAGATCAATCTGGGAATGAGGCAAGAAATTTTCAATGCGGCAAGGGGGCCCAT 1200
1309 TTTCCAGATCAATCTGGGAATGAGGCAAGAAATTTTCAATGCGGCAAGGGGGCCCAT 1368
1201 AGTGCACACCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
1369 AGTGCACACCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1428
1261 GAAATGATCACTGTCAACACTGTGATGATGATGATGATGATGATGATGATGATGAT 1320
1429 GAAATGATCACTGTCAACACTGTGATGATGATGATGATGATGATGATGATGATGAT 1488
1321 ATATATGTGTGCGAGTACCATTC 1344
1489 ATATATGTGTGCGAGTACCATTC 1512

RESULT 4
LOCUS BC022280
DEFINITION Homo sapiens, clone MGC:22412 IMAGE:4693953, mRNA, complete cds.
ACCESSION BC022280
VERSION BC022280.1 GI:18490144
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2367)
Strausberg, R.
Submitted (01-FEB-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: gcaps-remail.nih.gov
Tissue Procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.

CDs

1. .1347
 /gene="FIBL-5"
 /codon_start=1
 /product="Fibulin-5"
 /protein_id="CAB38568.1"
 /db_xref="GI:4490530"
 /translation="MPEIKRILTYTLALCLPSPNAOAOCTNGEDFDLROSGGOCID
 ECTIPEDRCMDVONNGVYLCIPRTNPPYRGPYSPYSGPYPAPAPLAP
 NYPTISPLICRFQYQDNESNOCVDDECATDSDHOCNPOTICITREGSTCTDGY
 LLEGQCLDIDECRYGCOQLCANVPSSYSCCTNPGFTLINEDRSCODVNECATENPCV
 QCVNTVGSFICRCDPGELEEDGVHCSDECSFSEFLNEHCNVOPGTFCSPG
 YLLDNRSODINEEHRNHTCNLOQTNYNLQGFKCIDIRCEEPYLRISDNCMC
 PANPCRDOPFTILYRDMDVYSRGPADIFOMATRYGAYIYIPIKSNMCMC
 YMRNGPISATLVMPRIKGRPRIDQLDMITVNTVINFRRSVIRLRITYSQIP"

sig_peptide
 1. .66
 /gene="FIBL-5"
 /gene="FIBL-5"
 /product="Fibulin-5"

mat_peptide
 67. .1344
 /gene="FIBL-5"
 /product="Fibulin-5"

BASE COUNT 507 a 540 c 475 g 496 t 1 others

ORIGIN

Query Match 100.0%; Score 1344; DB 9; Length 2019;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGCCAGGAATAAAGATACCTACCTGTTACCATTTCTGGCTCTCTCTTCCAAAGCCCT 60
 1 ATGCCAGGAATAAAGATACCTACCTGTTACCATTTCTGGCTCTCTCTTCCAAAGCCCT 60
 61 GGGAAATGCACAGGACAGTGCAGAAATGCTTTGACCTGATCCGACATCAGACAGTGT 120
 61 GGGAAATGCACAGGACAGTGCAGAAATGCTTTGACCTGATCCGACATCAGACAGTGT 120
 61 GGGAAATGCACAGGACAGTGCAGAAATGCTTTGACCTGATCCGACATCAGACAGTGT 120
 121 TTAGATATTGATGATATGCCGACCATCCCGAGGCTTCCGAGAGAGATGATGTGT 180
 121 TTAGATATTGATGATATGCCGACCATCCCGAGGCTTCCGAGAGAGATGATGTGT 180
 121 TTAGATATTGATGATATGCCGACCATCCCGAGGCTTCCGAGAGAGATGATGTGT 180
 181 AACCAAAATGGGGGATTTATGATTCCTCCGAGCAAAACCTGTGTATCCAGGGCCCTAC 240
 181 AACCAAAATGGGGGATTTATGATTCCTCCGAGCAAAACCTGTGTATCCAGGGCCCTAC 240
 181 AACCAAAATGGGGGATTTATGATTCCTCCGAGCAAAACCTGTGTATCCAGGGCCCTAC 240
 241 TCGAACCCCTACTCGACCCCTACTCAGTCCGTACCCAGCAGAGCTGCCACACACTCTCA 300
 241 TCGAACCCCTACTCGACCCCTACTCAGTCCGTACCCAGCAGAGCTGCCACACACTCTCA 300
 241 TCGAACCCCTACTCGACCCCTACTCAGTCCGTACCCAGCAGAGCTGCCACACACTCTCA 300
 301 GCTCCAAATATCCCAAGATCTCCAGGCTCTTATATGCCGCTTGGATACCAAGATGAT 360
 301 GCTCCAAATATCCCAAGATCTCCAGGCTCTTATATGCCGCTTGGATACCAAGATGAT 360
 301 GCTCCAAATATCCCAAGATCTCCAGGCTCTTATATGCCGCTTGGATACCAAGATGAT 360
 361 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAAAAGATTCGCCCAAGTGAACCCC 420
 361 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAAAAGATTCGCCCAAGTGAACCCC 420
 361 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAAAAGATTCGCCCAAGTGAACCCC 420
 421 ACCGATCTGATCAATACTGTAAGGCGGGGTACACCTGTCTCTGACCGAGATATTTGG 480
 421 ACCGATCTGATCAATACTGTAAGGCGGGGTACACCTGTCTCTGACCGAGATATTTGG 480
 421 ACCGATCTGATCAATACTGTAAGGCGGGGTACACCTGTCTCTGACCGAGATATTTGG 480
 421 ACCGATCTGATCAATACTGTAAGGCGGGGTACACCTGTCTCTGACCGAGATATTTGG 480
 481 CTTCTGGAAGGCGCATGCTTAGACATGTAATGTCTGATGTGTTACTGCCACAGCTC 540
 481 CTTCTGGAAGGCGCATGCTTAGACATGTAATGTCTGATGTGTTACTGCCACAGCTC 540
 481 CTTCTGGAAGGCGCATGCTTAGACATGTAATGTCTGATGTGTTACTGCCACAGCTC 540
 481 CTTCTGGAAGGCGCATGCTTAGACATGTAATGTCTGATGTGTTACTGCCACAGCTC 540
 541 TGTGGAATGTCTCTGATCTATCTGTACATGTAACCTGTTTACCCTCAATGAG 600
 541 TGTGGAATGTCTCTGATCTATCTGTACATGTAACCTGTTTACCCTCAATGAG 600
 541 TGTGGAATGTCTCTGATCTATCTGTACATGTAACCTGTTTACCCTCAATGAG 600
 541 TGTGGAATGTCTCTGATCTATCTGTACATGTAACCTGTTTACCCTCAATGAG 600
 601 GATGGAAGGTCTTGGCAAGATGTGAAGAGTGTGCCACGAGAACCCCTCGTGCACAAAC 660
 601 GATGGAAGGTCTTGGCAAGATGTGAAGAGTGTGCCACGAGAACCCCTCGTGCACAAAC 660
 601 GATGGAAGGTCTTGGCAAGATGTGAAGAGTGTGCCACGAGAACCCCTCGTGCACAAAC 660
 601 GATGGAAGGTCTTGGCAAGATGTGAAGAGTGTGCCACGAGAACCCCTCGTGCACAAAC 660
 661 TGGCTCAACACCTACGCGCTCTTCACTCTCCGCTGTGACCCAGATATGAATTTGAGAA 720
 661 TGGCTCAACACCTACGCGCTCTTCACTCTCCGCTGTGACCCAGATATGAATTTGAGAA 720
 661 TGGCTCAACACCTACGCGCTCTTCACTCTCCGCTGTGACCCAGATATGAATTTGAGAA 720
 661 TGGCTCAACACCTACGCGCTCTTCACTCTCCGCTGTGACCCAGATATGAATTTGAGAA 720

QY 721 GATGGCTTCATTGCACTGATATGAGAGAGTGCAGCTTCTCTGAGTTCCTCTGCCAAT 780
 DB 721 GATGGCTTCATTGCACTGATATGAGAGAGTGCAGCTTCTCTGAGTTCCTCTGCCAAT 780
 QY 781 GATGTGTGAACACAGCCGCGCACATCTTCTGCTCTCTGCTCCAGGCTACATCTGCTG 840
 DB 781 GATGTGTGAACACAGCCGCGCACATCTTCTGCTCTCTGCTCCAGGCTACATCTGCTG 840
 QY 841 GATGACAACCGAAGCTCCCAAGACATCAAGATGTGAGCAGAGAACACACAGTGCAC 900
 DB 841 GATGACAACCGAAGCTCCCAAGACATCAAGATGTGAGCAGAGAACACACAGTGCAC 900
 QY 901 CTGACAGACAGTGTCAATTTTCAAGAGGGGCTTCAAAATGATGACCCCATCCGCTGT 960
 DB 901 CTGACAGACAGTGTCAATTTTCAAGAGGGGCTTCAAAATGATGACCCCATCCGCTGT 960
 QY 961 GAGGAGCCTTATGTGAGATGATGATACCCGCTGTATGTCTCTGCTGAGAACCCCTG 1020
 DB 961 GAGGAGCCTTATGTGAGATGATGATACCCGCTGTATGTCTCTGCTGAGAACCCCTG 1020
 QY 1021 TGCAGAGACAGCCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
 DB 1021 TGCAGAGACAGCCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
 QY 1081 GTTCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
 DB 1081 GTTCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
 QY 1141 TTCCAGATCAAAATCTGGGAATGAGGGGAGAGAAATTTTCAATGGGGGAGGGCCCT 1200
 DB 1141 TTCCAGATCAAAATCTGGGAATGAGGGGAGAGAAATTTTCAATGGGGGAGGGCCCT 1200
 QY 1201 ACTGCCACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
 DB 1201 ACTGCCACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
 QY 1261 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
 DB 1261 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
 QY 1321 ATATATGTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1344
 DB 1321 ATATATGTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1344

RESULT 3
 AF112152 2328 bp mRNA linear PRI 02-AUG-1999
 LOCUS AF112152
 DEFINITION Homo sapiens developmental arteries and neural crest EGF-like
 protein mRNA, complete cds.
 ACCESSION AF112152
 VERSION AF112152.1 GI:5305672
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 (bases 1 to 2328)
 Nakamura,T., Kohno,K., Takahara,K., Ito,Y., Ando,N., Ito,S., Schaub,R.,
 Nakamura,A., Saegusa,S., Giken,K., and Honjo,Y.
 DAB2, a novel secreted EGF protein expressed in developing,
 embryonic, and balloon-injured arteries.
 JOURNAL OF BIOLOGICAL CHEMISTRY 274 (32): 22476-22483 (1999)
 MEDLINE 99357479
 PUBMED 10428823
 REFERENCE 2 (bases 1 to 2328)
 Nakamura,T., Yabe,D., Tashiro,K. and Honjo,T.
 Direct Submission
 Submitted (09-DEC-1998) Medicine, University of California San
 Diego, 9500 Gilman Drive, La Jolla, CA 92093-0613, USA
 FEATURES
 Location/Qualifiers

FEATURES
source

CDS

Location/Qualifiers
1..1720
/organism="unidentified"
/db_xref="taxon:32644"
211..1557
/note="unnamed protein product"
/protein_id="CAB69285.1"
/db_xref="GI:6733225"
/translation="MGKIRILVTILALCLPSRMAOACTNPDLDROSGCIDI
EERTIREACRGMACVNONGGLICIPRTNRYRGRYSPSTPSPYPAAPPLSAP
MPTISRLICRFGRYOMDESNOACVDECATDSHQNPOTICINTEGGYTSCTGYW
LLEQCLDIDECRYGCOQLCANVPSYSCNPNPFLNEDGRSDVNECATPCV
QICVNTYGSFICRCDPGLYELBDCGSDMDECSFSEFLNEDGRSDVNECATPCV
YLLDNRSQDINECEHRNHTCLQOTCYNLQSGFKIDIEIRYCEPLRISDNRCM
PAPNPGCRDPFTILYRDMDVSGRSVPADIFOMATTRYRAYVYIPIKSGNEGREF
VNRQIPISATLVMTPIKIPREIQLDLEMTVTNVINFRSSVIRLAIYVSQYF"

sig_peptide 211..284
mat_peptide 287..1554
BASE COUNT 396 a 515 c 434 g 372 t 3 others
ORIGIN

Query Match 100.0%; Score 1344; DB 6; Length 1720;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGCCAGATAAAGATGATCTACATTCCTGCTCTCTTCCAGCCCT 60
|||||
211 ATGCCAGATAAAGATGATCTACATTCCTGCTCTCTTCCAGCCCT 270
61 GGGATGACAGGACAGTGCAGGAATGGCTTTGACCTGATCGCCAGTCAAGAGTGT 120
|||||
271 GGGATGACAGGACAGTGCAGGAATGGCTTTGACCTGATCGCCAGTCAAGAGTGT 330
121 TTAGATATGATGATAATGCCAAGCATCCCGAGGCTGCCAGAGAGATATGTGTGT 180
|||||
331 TTAGATATGATGATAATGCCAAGCATCCCGAGGCTGCCAGAGAGATATGTGTGT 390
181 AACCAAAATGGCGGGATTATGATTCCTCCGGACAAACCTGTGTATGAGAGGCTCTAC 240
|||||
391 AACCAAAATGGCGGGATTATGATTCCTCCGGACAAACCTGTGTATGAGAGGCTCTAC 450
241 TGGAAACCTCTACCTGACCCCTCTACTCAGGTCCGTAACCAAGCTGCCCCACCACTCTCA 300
|||||
451 TGGAAACCTCTACCTGACCCCTCTACTCAGGTCCGTAACCAAGCTGCCCCACCACTCTCA 510
301 GGTCCAACATGACCAAGATCTCCAGGCGCTCTTATATGCGCTTGGATACAGATGAT 360
|||||
511 GGTCCAACATGACCAAGATCTCCAGGCGCTCTTATATGCGCTTGGATACAGATGAT 570
361 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCACACAGATTCACCACAGTGCACACCC 420
|||||
571 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCACACAGATTCACCACAGTGCACACCC 630
421 ACCCAATCTGCATCAATTAAGAGGGGGTACACCTGCTCTGACACGAGCGGATATGG 480
|||||
631 ACCCAATCTGCATCAATTAAGAGGGGGTACACCTGCTCTGACACGAGCGGATATGG 690
481 CTTCTGAAGGCGAGTCTAGACATTTGATGAATGTGCTATGTTTACTGCGCAGAGCTTC 540
|||||
691 CTTCTGAAGGCGAGTCTAGACATTTGATGAATGTGCTATGTTTACTGCGCAGAGCTTC 750
541 TGTGCAATGTTCTGTGATCTTCTTGTGATGACCAACCTGTGTTTACCCTCATATGAG 600
|||||
751 TGTGCAATGTTCTGTGATCTTCTTGTGATGACCAACCTGTGTTTACCCTCATATGAG 810
601 GATGGAAGTCTTGTGCAAGATGTGACAGAGTGTGCCACGAGAACCCCTGCGTGAACAC 660
|||||
811 GATGGAAGTCTTGTGCAAGATGTGACAGAGTGTGCCACGAGAACCCCTGCGTGAACAC 870
661 TGGGTCAACCACTACGGCTCTTTCATCTGCGGCTGTGACCAAGATATGAATGAGGAA 720
|||||
871 TGGGTCAACCACTACGGCTCTTTCATCTGCGGCTGTGACCAAGATATGAATGAGGAA 930

QY 721 GATGCGCTTCATGTCAGATATGAGAGTGCACGCTTCTGAGTTCCTTGCCAAACAT 780
|||||
DB 931 GATGCGCTTCATGTCAGATATGAGAGTGCACGCTTCTGAGTTCCTTGCCAAACAT 990
QY 781 GAGTGTGTAACCAAGCCCGGACATCTTCTGCTCTGCTCCACAGGCTACATCTGCTG 840
|||||
DB 991 GAGTGTGTAACCAAGCCCGGACATCTTCTGCTCTGCTCCACAGGCTACATCTGCTG 1050
QY 841 GATGCAACCAAGCACTGCGCAAGATCAACGAATGTGACAGAGAACCAACAGCTCAAC 900
|||||
DB 1051 GATGCAACCAAGCACTGCGCAAGATCAACGAATGTGACAGAGAACCAACAGCTCAAC 1110
QY 901 CTGACAGAGAGCTGCTCAATTTCAAGAGGCTTCAATGATGATGACACCCCAATCCCTGT 960
|||||
DB 1111 CTGACAGAGAGCTGCTCAATTTCAAGAGGCTTCAATGATGATGACACCCCAATCCCTGT 1170
QY 961 GAGAGCCCTTATCTGAGGATGATGATTAACGCTTATGTCTGCTGTAAGAACCTGTGC 1020
|||||
DB 1171 GAGAGCCCTTATCTGAGGATGATGATTAACGCTTATGTCTGCTGTAAGAACCTGTGC 1230
QY 1021 TGCAGAGACCAAGCCCTTATACATCTTGTACCGGAGATGAGAGCTGTGTCAGAGCCTCC 1080
|||||
DB 1231 TGCAGAGACCAAGCCCTTATACATCTTGTACCGGAGATGAGAGCTGTGTCAGAGCCTCC 1290
QY 1081 GTTCCCGCTGACATCTTCCAAATGCAAGCAGACGACCCGCTACCTGCGGCTTATACAT 1140
|||||
DB 1291 GTTCCCGCTGACATCTTCCAAATGCAAGCAGACGACCCGCTACCTGCGGCTTATACAT 1350
QY 1141 TTCCAGATCAATCTGGAATGAGGCGAGAGCAATTTTACATGCGGCAACGGGCCCCATC 1200
|||||
DB 1351 TTCCAGATCAATCTGGAATGAGGCGAGAGCAATTTTACATGCGGCAACGGGCCCCATC 1410
QY 1201 AGTCCCAACCCGTGATGATGACAGCCGCCCATCAAGAGGCCCCGGGAATTCAGCTGACTTG 1260
|||||
DB 1411 AGTCCCAACCCGTGATGATGACAGCCGCCCATCAAGAGGCCCCGGGAATTCAGCTGACTTG 1470
QY 1261 GAAATGATCACTGTCAACACCTGTCTCAACTTCAGAGGAGCTCCGTGATTCGACGTGCGG 1320
|||||
DB 1471 GAAATGATCACTGTCAACACCTGTCTCAACTTCAGAGGAGCTCCGTGATTCGACGTGCGG 1530
QY 1321 ATATATGTGTGCGAGTACCAATTC 1344
|||||
DB 1531 ATATATGTGTGCGAGTACCAATTC 1554

RESULT 2
HSA133490
LOCUS HSA133490 2019 bp mRNA linear PRI 11-MAR-1999
DEFINITION Homo sapiens fibulin-5.
ACCESSION AJ133490
VERSION AJ133490.1 GI:4490529
KEYWORDS FIBL-5 gene; fibulin-5.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2019)
AUTHORS Kostka G.
JOURNAL Direct Submission
Submitted (05-MAR-1999) Kostka G., Dept. of Protein Chemistry,
Max-Planck-Institut fuer Biochemie, Am Klopferspitz, 82152
Martinsried, GERMANY
2 (bases 1 to 2019)
AUTHORS Kostka G.
JOURNAL Unpublished
FEATURES
source
Location/Qualifiers
1..2019
/organism="Homo sapiens"
/db_xref="taxon:9606"
/issue_type="melanoma"
1..1347
/gene="FIBL-5"

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 08:43:26 ; Search time 3504.57 Seconds

(without alignments)
11160.894 Million cell updates/sec

File: US-09-674-379a-11

Sequence: 1 atgcgcaggaataaaagat.....atgtctgcagtaaccattc 1344

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: GenBank
1: gb-bank
2: gb-bank
3: gb-in
4: gb-om
5: gb-ov
6: gb-pat
7: gb-ph
8: gb-pl
9: gb-pr
10: gb-ro
11: gb-sts
12: gb-sy
13: gb-un
14: gb-vl
15: em-ba
16: em-fun
17: em-hum
18: em-in
19: em-mu
20: em-om
21: em-or
22: em-ov
23: em-pat
24: em-ph
25: em-pl
26: em-ro
27: em-sts
28: em-un
29: em-vl
30: em-hg-hum
31: em-hg-inv
32: em-hg-other
33: em-hg-mus
34: em-hg-pln
35: em-hg-rod
36: em-hg-mam
37: em-hg-vrt
38: em-sy
39: em-hggo-hum
40: em-hggo-mus
41: em-hggo-other

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1344	100.0	1720	6 A84086	A84086 Sequence 1
2	1344	100.0	2019	9 HSA133490	AJ133490 Homo sapi
3	1344	100.0	2338	9 AF112152	AF112152 Homo sapi
4	1344	100.0	2367	9 BC022280	BC022280 Homo sapi
5	1344	100.0	2550	6 AR036548	AR036548 Homo sapi
6	1344	100.0	2550	6 AR173204	AR173204 Sequence
7	1342.4	99.9	2609	6 AX376340	AX376340 Sequence
8	1342.4	99.9	2609	6 AX403659	AX403659 Sequence
9	1338	98.8	2019	9 AF093118	AF093118 Homo sapi
10	1113.6	82.9	2230	10 BC006636	BC006636 Mus muscu
11	1113.6	82.9	2478	10 AF112151	AF112151 Mus muscu
12	1104	82.1	2499	10 AF112153	AF112153 Rattus no
13	1102.4	82.0	2304	10 AF137350	AF137350 Rattus no
14	372.8	27.7	1440	9 HSA132819	AJ132819 Homo sapi
15	372.8	27.7	1536	9 AB030655	AB030655 Homo sapi
16	371.2	27.6	1480	6 AX023967	AX023967 Sequence
17	371.2	27.6	1561	6 BC010456	BC010456 Homo sapi
18	371.2	27.6	1707	9 AK000980	AK000980 Homo sapi
19	371.2	27.6	1875	6 AX201326	AX201326 Sequence
20	371.2	27.6	1875	6 AX464184	AX464184 Sequence
21	371.2	27.5	2018	6 AR095382	AR095382 Sequence
22	369.6	27.5	1757	9 AF109121	AF109121 Homo sapi
23	364.8	27.1	1825	9 AF124486	AF124486 Homo sapi
24	363.8	27.1	1705	9 AF093119	AF093119 Homo sapi
25	350.4	26.1	1778	10 AF046870	AF046870 Cricetus
26	347.2	25.8	1512	10 AF104223	AF104223 Mus muscu
27	347.2	25.8	1513	6 AX023961	AX023961 Sequence
28	347.2	25.8	1794	10 AF109122	AF109122 Mus muscu
29	345.6	25.7	1724	10 BC012269	BC012269 Mus muscu
30	337.6	25.1	1358	6 AX023976	AX023976 Sequence
31	312	23.2	1021	6 AX023954	AX023954 Sequence
32	293.8	22.0	2613	9 AK098186	AK098186 Homo sapi
33	268	19.9	1122	6 AX023965	AX023965 Sequence
34	260.6	19.4	171901	9 AC007057	AC007057 Homo sapi
35	260.6	19.4	172674	9 CENS0000R	AL049872 Human chr
36	260.6	19.4	175210	9 CENS0000R	AL590328 Human chr
37	250.6	18.6	1677	10 BC031184	BC031184 Mus muscu
38	249	18.5	1911	10 DB9730	DB9730 Rattus rat
39	241	17.9	2025	9 BC014410	BC014410 Homo sapi
40	241	17.9	2512	6 AX277602	AX277602 Sequence
41	241	17.9	2512	6 AX331916	AX331916 Sequence
42	241	17.9	2512	6 AX332783	AX332783 Sequence
43	241	17.9	2512	6 AX333462	AX333462 Sequence
44	241	17.9	2512	6 AX334103	AX334103 Sequence
45	241	17.9	2512	6 AX335899	AX335899 Sequence

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL
A84086	A84086	Sequence 1 from Patent WO9846746.	A84086	A84086.1	GI:6733224	unclassified.	unclassified.	1 (bases 1 to 1720)	Li, H. and Olsen, H. S.	EXTRACELLULAR/EPIDERMAL GROWTH FACTOR LIKE PROTEIN	Reference WO 9846746-2 A 2 22-06-1998
						unclassified.	unclassified.				HUMAN GENOME SCIENCES INC (US) R. L. HAO DONG (US)


```

; ADDRESSEE: Williams, Morgan & Amerson
; STREET: 7676 Hillmont, Suite 250
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77040
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,722B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US95/02251
; FILING DATE: 21-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/316,650
; FILING DATE: 30-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/199,780
; FILING DATE: 18-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fussey, Shelley P.M.
; REGISTRATION NUMBER: 39,458
; REFERENCE/DOCKET NUMBER: 4100.000500/FUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 934-7000
; TELEFAX: (713) 934-7011
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3759 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3759
;
; US-08-479-722B-3
;
; Query Match
; Best Local Similarity 4.0%; Score 53.2; DB 3; Length 3759;
; Matches 167; Conservative 0; Mismatches 168; Indels 3; Gaps 1;
;
; QY 541 TGTGCGATGTCCTGGATCTTATCTTGTGATGCAACCCCTGTTTACCTCAATGAG 600
; DB 2215 TGTGAGAACCTTCCGGGCTCTTACCGTTCGACGTTGCCCCAGGGGATACGAACCCGACA 2274
;
; QY 601 GATGAGAGCTTGGCAAGATGAGACGATGAGCCACCGAAGACCCCTGC--GTGCAA 657
; DB 2275 GGACCCCTCAGTTCATTAACGGATGATGAGCGCTGGAAAGTGTGCCAAGATGGC 2334
;
; QY 658 ACCCTGCTCAACACCTACCTCTTTCATCTGCCGCTGTGACCCAGATATGAATGAG 717
; DB 2335 ATCTCAGAGAACACACCGAGCTCTTCCAGTGTGACGCTCTCCGCTATCACTGTGCA 2394
;
; QY 718 GAAGATGCGCTTCATTCAGTATGATGAGACGATGCAACCTTCTGTGATTCCTGTGCAA 777
; DB 2395 AGGAGTCGAGAGCGCTGTGAGGACATGATGATGATGATTCCTCGGCGCTGATCGGG 2454
;
; QY 778 CATGATGTGTATACCAACCGCCGACATATCTTGTCTCCCTGCGCTCCAGGCTAATCTG 837
; DB 2455 GGTGATCTGATTAATACCAATGATGATGATGATGATGATGATGATGATGATGATG 2514
;
; QY 838 CTGATGACACCGAAGCTGCGACATCAACGAATG 875
; DB 2515 GTGGCGGACGAGATGCAAGAAAGATATGATGATG 2552
;
; RESULT 13
; US-08-479-722B-1
```

```

; Sequence 1, Application US/08479722B
; Patent No. 6074840
; GENERAL INFORMATION:
; APPLICANT: Bonadio, Jeffrey
; APPLICANT: Yin, Wushan
; TITLE OF INVENTION: LATENT TGF ( BINDING PROTEIN (LTBP)
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Williams, Morgan & Amerson
; STREET: 7676 Hillmont, Suite 250
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77040
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,722B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US95/02251
; FILING DATE: 21-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/316,650
; FILING DATE: 30-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/199,780
; FILING DATE: 18-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fussey, Shelley P.M.
; REGISTRATION NUMBER: 39,458
; REFERENCE/DOCKET NUMBER: 4100.000500/FUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 934-7000
; TELEFAX: (713) 934-7011
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5499 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: /desc = "DNA"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..5499
;
; US-08-479-722B-1
;
; Query Match
; Best Local Similarity 3.9%; Score 52.4; DB 3; Length 5499;
; Matches 223; Conservative 0; Mismatches 241; Indels 6; Gaps 2;
;
; QY 556 GATCTCTATCTTGTATACATCAACCTGTTTACCTCAATGAGATGAGATGAGTCTTGC 615
; DB 3121 GCTCTCTTATGATGCTCTCTGTGAGCGGCGCTATGAGGTCAACCCAGACAAAGAGGCTGC 3180
;
; QY 616 CAAGATGAGAGAGTGTGACCGACCGAAGACCCCTG---CGTGCAAACTGTGTCACAC 672
; DB 3181 CGAGATGTGAGAGAGTGTGACCGACCGAGCTGTGCCCCAGGGGCTGTGCTCAACACG 3240
;
; QY 673 TAGCGCTCTTCACTGTGCGG---CTGTGACCCAGATATGAACTTGAGAGAGATGAGCTT 729
; DB 3241 GAGGCGCTCTTCACTGTGCTGACGCTGTGACAGCGGATGAGGATGAGAGAGAGGACT 3300
;
; QY 730 CATTCAGATGATGAGACGATGCAAGCTTCTGTGATGATGATGATGATGATGATGATG 789
; DB 3301 GCCTGTGAAGACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3360
```

```

Db      93 AGATCCCTACATCTTAACACACAGAAACCGATGCTGTTGCCAGCTCTCAATGCCATGTG 152
OY      1023 CAGAGACACCCCTTTTACATCTTGTACCGGACATGAGCGTGTGACGAGCCGCT 1082
Db      153 CCGAGAACTCCCGACGATAGTCTACAAATATACATGACATCCATCGATGATGGTCTGT 212
OY      1083 TCCCGGTGACATCTTCCAAATGACAGCAGCCGCTACCTGGGCTTATACATTTT 1142
Db      213 GCCATCAGACATCTTCCAGATACAGGCCACAACTATTTATGCAACACATCAATACCTTT 272
OY      1143 CCAGTAAATCTGCGGATGAGGCGACAGAAATTTTACATGCGGCAAGCGGCCCTACG 1202
Db      273 TCGGATTAAATCTGGAATGAAATGAGG-AGTCTACCTACGACAAACAAACCTCTTAAG 331
OY      1203 TGCACCCCTGCTGATGACACGCCCATCAAGGCGCCGGAATCCAGCTGGACTTGA 1262
Db      332 TGCATATCTTGTCTGCTGTAAGNCAATTTATCAGGACCAAGAACATATCTGACTGGA 391
OY      1263 AATGATCAC 1271
Db      392 GATGCTGAC 400

```

```

RESULT 10
US-09-404-879A-340/c
; Sequence 340, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 340
; LENGTH: 220
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(220)
; OTHER INFORMATION: n - A,T,C or G
US-09-404-879A-340

```

```

Query Match      6.0%; Score 80.2; DB 4; Length 220;
Best Local Similarity 60.5%; Pred. No. 5.5e-16;
Matches 130; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
OY      373 TGTGTGATGTGACGAGTGTGCAACAGATTCCACCAAGTGCACACCCACAGATCTGC 432
Db      219 TGTGTGATGTGACGAGTGTGCGCCAGCCCTGCGACGCTGTGCGCCACGAGGACTGC 160
OY      433 ATCAATAGTAAAGGCGGTACACCTGCTCACTGCAACCGAGATATTGGCTTGTGAAGGC 492
Db      159 CATAACTGCTGCTGCTCTTCACTGACATGCACTGCTGCTGTTACCGAAGATCGGGCC 100
OY      493 CAGTCTTACATGATGATGATGCTGCTATGTTACTGCGACAGACTCTGTGCGAATGTT 552
Db      99 GAGTGTGTGACATAGACGAGTGCCTGCTACCGCTACTGCCAGACCGGTGCTGAACCTG 40
OY      553 CCTGATCCATCTTCTTATACATGCAACCCGTGTTT 587
Db      39 CCGGCTCTTCCGCTGCGACGACNCCAGCCGGGCTT 5

```

```

RESULT 11
US-08-282-141-1
; Sequence 1, Application US/08282141
; Patent No. 5538861

```

```

; GENERAL INFORMATION:
; APPLICANT: Schneider, Claudio
; APPLICANT: Varnum, Brian
; APPLICANT: Avanzi, Giancarlo
; APPLICANT: Brancolini, Claudio
; APPLICANT: Manioletti, Guido
; TITLE OF INVENTION: Stimulating Factor for the AXL Receptor
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehaven Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: United States
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/282,141
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2461 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
US-08-282-141-1

```

```

Query Match      4.3%; Score 58.2; DB 1; Length 2461;
Best Local Similarity 54.5%; Pred. No. 2.7e-08;
Matches 164; Conservative 0; Mismatches 128; Indels 9; Gaps 2;

```

```

OY      521 ATGTTACTGCGACAGCTGTGCGAATGTTCTGATGATCTATCTGTACATGCAACC 580
Db      625 ACGGGGCTGCTCTCCAGATCTGCGCAACAAACCGGGTACCTTCTCTGCGACA 684
OY      581 CTGTTTACCTCAATGAGATGAGAGTCTTGCCAAAGATGACGAGTGTGCCACCG 640
Db      685 GCGCTTGACGCTCTCTCTGATGAGGAGACCTGCCAAGACATAGAGTGTGCGAGCT 744
OY      641 AGAACCCCTGCTGCAACCTGCGTCAACACCTA---CGGCTTTCACTGCGCTGTG 697
Db      745 CCGAGGCTGTGGGGAGCGCGCTGCAAGAACCTGCGGCTCTACTCTGCTGTG 804
OY      698 ACCGAGATATAGACTTGAGAGATGCGCTTCACTGACATATGACGAGTGCAGCT 757
Db      805 ACGAGGCTTGTGCTACAGCTCCACGAGAGAGGCTTGCCGAGATGTGACAGTG----- 859
OY      758 TCTGTGATTCCTTCCGACATGATGATGTGTGAACAGCCCGCACATCTTCTGCTCT 817
Db      860 -TCTGAGGGCGCGTGTGACGAGTCTCGTGAACCTCCACAGGAGCTACACTGCGACT 918
OY      818 G 818
Db      919 G 919

```

```

RESULT 12
US-08-479-722B-3
; Sequence 3, Application US/08479722B
; Patent No. 6074840
; GENERAL INFORMATION:
; APPLICANT: Bonadio, Jeffrey
; APPLICANT: Yin, Wushan
; TITLE OF INVENTION: LATENT TGF (BINDING PROTEIN (LTBP)
; TITLE OF INVENTION: GENES, COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:

```


APPLICANT: Onrust, Rene
APPLICANT: Murison, James Greg
TITLE OF INVENTION: Compositions Isolated From Skin Cells
FILE REFERENCE: 11000.1011C
CURRENT APPLICATION NUMBER: US/09/188.930A
CURRENT FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 348
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO: 259
LENGTH: 1018
TYPE: DNA
ORGANISM: Human
US-09-188-930-259

Query Match 23.68; Score 316.6; DB 3; Length 1018;
Best Local Similarity 62.38; Pred. No. 2.9e-91;
Matches 515; Conservative 0; Mismatches 309; Indels 3; Gaps 1;

518 GCTATGTTACTGTCAGCAGCTCTGTGAGATGTTCCGATCCATCTTGTACATGCA 577
1 GCTACCGCTACTGTCAGCAGCTCTGTGAGATGTTCCGATCCATCTTGTACATGCA 60
578 ACCCTGTTTACCTCAGTATGAGATGAGAGCTTGTCCAGATGTAAGAGTGTGCA 637
61 AGCCGGCTTCCAGCTGGGCTTAACAACCGCTCTGTGTGATGTAAGAGTGTGCA 120
638 CCGAGAACCCCTGTGAGAACCTGCGTCAACCTAGAGCTTGTCCAGATGTAAGAGTGTGCA 697
121 TGGGGGCCCCATGAGACAGCGCTGCTCAACCTCTATGAGACCTCTGTGTGCTGCG 180
698 ACCAGATATGAACTTGAAGATGAGAGCTTGTATGAGATGTAAGAGTGTGCA 757
181 ACCAGGCTATGAGCTCATGCGATGAGTGTCTCTCTGAGATGTAAGAGTGTGCA 240
758 TCTCTAGTCTCTGTGCAACATGATGTGTGAGACCGCGCAATCTTGTCTGCT 817
241 ACTCCAGCTACTCTGTGAGTACAGCTGCTGCAAGAGACCGCGCTTCTCTGCT 300
818 GCGCTCAGGCTACTCTGTGAGTACAGCTGCTGCAAGAGACCGCGCTTCTCTGCT 877
301 GCGCACAGGCTTACAGCTGCTG---GCCACAGCGCTTGTCCAGATGTAAGAGTGTG 357
878 AGCAGAGAACACACAGCTGCAACCTGAGACAGCTGCTTATTAAGAGGCGCTTCA 937
358 AGCTGCTGAGACAGCTGCTGCGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 417
938 AATGATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 997
418 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 477
998 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1057
478 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 537
1058 TGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1117
538 TGACCATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 597
1118 GCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1177
598 TCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 657
1178 ACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1237
658 ACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717
1238 CCGGAGAAATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1297
718 CCGGAGAAATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 777
1298 GCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1344

Db 778 CCAGCTCTGATGAGGCTCACCGCTTGTGAGGGGCTTACACCTTC 824
RESULT 7
US-09-248-757-1
Sequence 1, Application US/09248757
Patent No. 6417342
GENERAL INFORMATION:
APPLICANT: STONE, EDWIN M.
APPLICANT: SHEFFIELD, VAL C.
TITLE OF INVENTION: MACULAR DEGENERATION DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: UIA-018.02
CURRENT APPLICATION NUMBER: US/09/248.757
CURRENT FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 1
LENGTH: 2512
TYPE: DNA
ORGANISM: Homo sapiens
US-09-248-757-1

Query Match 17.98; Score 241; DB 4; Length 2512;
Best Local Similarity 54.68; Pred. No. 8.1e-67;
Matches 547; Conservative 0; Mismatches 445; Indels 9; Gaps 3;

346 GGATACAGATGATGAAAGCAACCAATGTGTGATGAGAGAGTGTGCAACAGATTCC 405
403 GCGTACAGCAAGTGAACACACAGCTGTGCCAAGATGAGAGAGTGTGCAACAGATTCC 462
406 CACCACTGCAACCCACACAGATGCTCATTAATGAGAGGCGGTACACCTGCTCTGC 465
463 CACAACGTGAGAGACAGCAAGATGCTCATTAATGAGAGGCGGTACACCTGCTCTGC 522
466 ACCAGAGATGTTGCTCTGTGAGAGCCAGTCTTATGATGATGATGATGATGATGAT 522
523 CCTCTGATGATGAG 582
523 GGTATGCTGAG 582
583 CCATATGCTGAG 642
643 AACCTGCTGAG 702
703 AATCAATGCTGAG 762
703 GGATATGATGAG 762
763 GGATATGATGAG 822
763 GAGTTCCTGCTGAG 822
823 AGTACCTGCTGAG 882
823 CAGGCTACATCTGCTGAG 882
883 CAGGATA---CCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 939
883 AGGAGACAG 942
940 ACAATATCA---ATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 996
943 ATGACCCCATCTGCTGAG 1002
997 TATCAGAGAAATCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1056
1003 CTTCTGAG 1062
1057 CCACTCTCAATGAG 1116

QY 1117 CGCTACCTGGGGCCTATTACATTTCCAGATCAATCTGGGAATGAGGCGACAGATTT 1176
Db 1226 GTCTACCCCGGCTCTACATGCTTTTCAGATCCGTGCGAATCTGCGAGGGGACTTT 1285
QY 1177 TACATGCGGCAAAAGGGCCCATCACTGATGACAGCCCATCAAAAGG 1236
Db 1286 TACATTAGGCAATCAACAACTGACGCGCATGTCTCTGCGCGCGGCGGTGAGAGGC 1345
QY 1227 CCGCGGGAATCCACTGACTTGGAAATGATCACTGTCAACACTGTCACTCAAG 1296
Db 1346 CCGCGGGAATCCACTGACTTGGAAATGATCACTGTCAACACTGTCACTCAAG 1405
QY 1297 GCGAGCTCCGTGATCCGATCGGAGATATGTGTGCGAGTACCATTC 1344
Db 1406 GCGAGCTCCGTGATCCGATCGGAGATATGTGTGCGAGTACCATTC 1453

RESULT 5

US-08-980-514-2
; Sequence 2, Application US/08980514
; Patent No. 6004753
; GENERAL INFORMATION:
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN SI-5-ECMP-LIKE PROT
; TITLE OF INVENTION: E1N
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/980,514
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0436 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2018 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRESTNOT13
; CLONE: 2786449
; US-08-980-514-2

Query Match 27.6%; Score 371.2; DB 3; Length 2018;
Best Local Similarity 61.2%; Pred. No 1.5e-108;
Matches 617; Conservative 0; Mismatches 388; Indels 3; Gaps 1;

QY 337 TCCCGCTTTGATACAGATGATGAAGCAACCAATGTGTGATGTGAGAGTGTGCA 396
Db 530 TCCCGCAGGCTATGAGCCCGAGCATGAGACAGCTGTGTGATGTGAGAGTGTGCC 589

QY 397 ACAGATTCCACCACTGCAACCCCAACAGATGTGCATCAATCTGAAAGCGGGTACACC 456
Db 590 CAGGCGCCGACAGACTGTGCGCCCGAGGAGCTGCATTAATCTGCGTGTCTATCAG 649
QY 457 TGCCTCCGACCGAGGATATTGGCTTCTGGAAGCGCATGCTTGAATGATGATGT 516
Db 650 TGCAGCTGCGCTGATGTTTACCGCAAGATCGGGCCGAGTGTGTGAGCATGACAGTGC 709
QY 517 CGCTATGTTTACTCCAGACGCTGTGTGATGATGTTTCTGATGATGATGATGATG 576
Db 710 CGCTACCGCTATGCTGACAGACCGCTGTGATGATGATGATGATGATGATGATG 769
QY 577 AACCTGTTTACCTCAATGAGAGTGAAGTCTTCCAGATGATGATGATGATGATG 636
Db 770 GAGCGGGGCTTCCAGCTGCGGCTTCAACACCGCTCTGTGTGATGATGATGATGATG 829
QY 637 ACCGAGAACCCCTGTCGCAACCTGCTCAACCTGCTTATCTGCTGCTGCTGCTG 696
Db 830 ATGGGGGCGCCATGCGAGACGCGCTGCTTCAACTCTGTGAGACCTTCTGCTGCTG 889
QY 697 GACCGAGATATGACTTGAAGAGATGCGCTTCACTGATGATGATGATGATGATG 756
Db 890 CACGAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 949
QY 757 TTTCTGATGTTCTCTGCTGCAATGATGATGATGATGATGATGATGATGATGATG 816
Db 950 TACTCAGCTTACTCTCTGATGATGATGATGATGATGATGATGATGATGATGATG 1009
QY 817 TGCCCTCCAGCTACATCTGCTGATGATGATGATGATGATGATGATGATGATGATG 876
Db 1010 TGCCCGACAGGTTTACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1066
QY 877 GAGCAGAGAACCAAGCTGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 936
Db 1067 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1126
QY 937 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 996
Db 1127 CGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1186
QY 997 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1056
Db 1187 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1246
QY 1057 ATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1116
Db 1247 ATGACCATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1306
QY 1117 CGCTACCTGGGGCCTATTACATTTCCAGATCAATCTGGGAATGAGGCGACAGATTT 1176
Db 1307 GTCTACCCCGGCTCTACATGCTTTTCAGATCCGTGCGAATCTGCGAGGGGACTTT 1366
QY 1177 TACATGCGGCAAAAGGGCCCATCACTGATGACAGCCCATCAAAAGG 1236
Db 1367 TACATTAGGCAATCAACAACTGACGCGCATGTCTCTGCGCGCGGCGGTGAGAGGC 1426
QY 1237 CCGCGGGAATCCACTGACTTGGAAATGATCACTGTCAACACTGTCACTCAAG 1296
Db 1427 CCGCGGGAATCCACTGACTTGGAAATGATCACTGTCAACACTGTCACTCAAG 1486
QY 1297 GCGAGCTCCGTGATCCGATCGGAGATATGTGTGCGAGTACCATTC 1344
Db 1487 GCGAGCTCCGTGATCCGATCGGAGATATGTGTGCGAGTACCATTC 1534

RESULT 6

US-09-188-930-259
; Sequence 259, Application US/09188930A

; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew

```

Db 367 ACCAKATNATGAGTGGAGAGATGGCATTCATGCAATATGATGATGACACT 426
Qy 758 TCTGTAGTTCCTCTGCCAATGATGTGTGACACCGCCGACATCTTCTGCTCT 817
Db 427 TCTCCGAGTTCCTCTGTCATGATGTGTGACACCGCCGACATCTTCTGCTCAT 486
Qy 818 GCCCTCAGGCTGATCTGCTGATGATGATGATGATGATGATGATGATGATGAT 877
Db 487 GCCCTCAGGCTGATCTGCTGATGATGATGATGATGATGATGATGATGATGAT 546
Qy 878 AGCAGAGAACACACAGTGCAGACCTGCAGACAGCTGCATTAATTAACAGGGGCTTCA 937
Db 547 AGCAGAGAACACACAGTGCAGACCTGCAGACAGCTGCATTAATTAACAGGGGCTTCA 606
Qy 938 AATGATATGACACCTGCTGCTGTGTGAGAGCTTATCTGATGATGATGATGATGAT 997
Db 607 AATGATATGACACCTGCTGCTGTGTGAGAGCTTATCTGATGATGATGATGATGAT 666
Qy 998 TGTGCTGCTGAGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1057
Db 667 TGTGCTGCTGAGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 726
Qy 1058 TGGAGCTGTGTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1117
Db 727 TGGAGCTGTGTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 786
Qy 1118 GCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1177
Db 787 GATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 846
Qy 1178 ACATGCGGCAAAAGGCGCTATGATGATGATGATGATGATGATGATGATGATGAT 1237
Db 847 ACATGCGGCAAAAGGCGCTATGATGATGATGATGATGATGATGATGATGATGAT 906
Qy 1238 CCGGGAATTCACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1297
Db 907 CCGGGAATTCACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 966
Qy 1298 GCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1344
Db 967 GCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1013

```

RESULT 4

US-08-833-963C-1
 ; Sequence 1, Application US/08833963C
 ; Patent No. 5916769

GENERAL INFORMATION:

APPLICANT: Olsen, et al.
 TITLE OF INVENTION: Extracellular/Epidermal Growth Factor
 TITLE OF INVENTION: HCBAB58X
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Ave
 CITY: Rockville
 STATE: MD
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/833,963C
 FILING DATE: 11-APR-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/US96/05033
 FILING DATE: 10-APR-1996
 ATTORNEY/AGENT INFORMATION:

```

NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PR258
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 125..1453
US-08-833-963C-1
Query Match 27.6%; Score 371.2; DB 2; Length 1531;
Best Local Similarity 61.2%; Pred. No. 1.3e-108;
Matches 617; Conservative 0; Mismatches 388; Indels 3; Gaps 1;
Qy 337 TGGCGCTTTGGATACCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 396
Db 449 TGGCGCGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 508
Qy 397 ACAGATTCACACAGTGCAGACCCGACAGATGATGATGATGATGATGATGATGATGAT 456
Db 509 CAGGCGCTGCAGAGTGTGCGCCGACGAGAGTGCATTAATGCTGCTGCTGCTGCTGCT 568
Qy 457 TGTCTCTGACCGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 516
Db 569 TGCACCTGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 628
Qy 517 CGCTATGCTTACTGTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 576
Db 629 CGCTATGCTTACTGTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 688
Qy 577 AACCTGCTTACTGTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 636
Db 689 GAGCGGCTTCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 748
Qy 637 ACCGAGAACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 696
Db 749 ATGGGGGCGCCATGCGAGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 808
Qy 697 GACCCAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 756
Db 809 CACGAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 868
Qy 757 TTTCTGAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 816
Db 869 TACTTCAGCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 928
Qy 817 TGCCTCAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 876
Db 929 TGCCTCAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 985
Qy 877 GAGCAGAGAACACACAGTGCAGACCTGCAGACAGCTGCTGCTGCTGCTGCTGCTGCT 936
Db 986 GAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1045
Qy 937 AATGATATGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 996
Db 1046 CGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1105
Qy 997 ATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1056
Db 1106 CTGTCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1165
Qy 1057 ATGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1116
Db 1166 ATGACATTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1225

```

```
OY 1 ATGCCAGGAATATAAAGGATACCTGATACATTTCTGCTCTCTCTCTCCAGCCCT 60
DB 382 ATGCCAGGAATATAAAGGATACCTGATACATTTCTGCTCTCTCTCTCCAGCCCT 441
OY 61 GGGATATCAGACGACAGTGCAGCAATGGCTTTGACCTGAGTGGCAGTACAGAGTGT 120
DB 442 GGGATATCAGACGACAGTGCAGCAATGGCTTTGACCTGAGTGGCAGTACAGAGTGT 501
OY 121 TTGATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
DB 502 TTGATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 561
OY 181 AACCAAAATGGGGGATTTTATGATTTCCCGGAGAACCCCTGTATGAGAGGCTTAC 240
DB 562 AACCAAAATGGGGGATTTTATGATTTCCCGGAGAACCCCTGTATGAGAGGCTTAC 621
OY 241 TCGAACCCCTACTGACGACCCCTACTGAGTCCGTACCCAGAGAGTGGCCACACTTCA 300
DB 622 TCGAACCCCTACTGACGACCCCTACTGAGTCCGTACCCAGAGAGTGGCCACACTTCA 681
OY 301 GCTCCAACTATCCAGATCTCCAGGCTCTTATATGCTGCTTGTATACCAATGAT 360
DB 682 GCTCCAACTATCCAGATCTCCAGGCTCTTATATGCTGCTTGTATACCAATGAT 741
OY 361 GAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAACAGATTTCCACAGTGCACCC 420
DB 742 GAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAACAGATTTCCACAGTGCACCC 801
OY 421 ACCCAGATCTGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 802 ACCCAGATCTGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 861
OY 481 CTCTCTGGAAGGCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 862 CTCTCTGGAAGGCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 921
OY 541 TGTGCAATGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 922 TGTGCAATGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 981
OY 601 GATGAGAGGTCTTGCAGATGTGAGAGAGTGTGCAACGAGAACCCCTGCTGCAAC 660
DB 982 GATGAGAGGTCTTGCAGATGTGAGAGAGTGTGCAACGAGAACCCCTGCTGCAAC 1041
OY 661 TGGGTAAACCAAGGCTCTTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 1042 TGGGTAAACCAAGGCTCTTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1101
OY 721 GATGCGTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 1102 GATGCGTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1161
OY 781 GAGTGTGTGACAGCCCGGACATATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
DB 1162 GAGTGTGTGACAGCCCGGACATATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1221
OY 841 GATGACAAACGAGAGTGTGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 1222 GATGACAAACGAGAGTGTGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1281
OY 901 CTGACAGACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 1282 CTGACAGACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1341
OY 961 GAGGAGCCCTTATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 1342 GAGGAGCCCTTATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1401
OY 1021 TGCAGAGACAGCCCTTATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1402 TGCAGAGACAGCCCTTATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1461
```

```
OY 1081 GTTCCCGCTGACATCTTCCAAATGACAGCCAGCCGCTACCTGGGGCTATTATATT 1140
DB 1462 GTTCCCGCTGACATCTTCCAAATGACAGCCAGCCGCTACCTGGGGCTATTATATT 1521
OY 1141 TTCCAGATCAATGTGGGAATGAGGGGAGAGATTTTACATGCGGCAAGGGGCCCATC 1200
DB 1522 TTCCAGATCAATGTGGGAATGAGGGGAGAGATTTTACATGCGGCAAGGGGCCCATC 1581
OY 1201 AGTGCACCCCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
DB 1582 AGTGCACCCCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1641
OY 1261 GAATGATCACTGTCAACACCTGTATCACTTACAGAGCAGCTCGTATCCGACTGGG 1320
DB 1642 GAATGATCACTGTCAACACCTGTATCACTTACAGAGCAGCTCGTATCCGACTGGG 1701
OY 1321 ATATATGTGCGAGTACCCATTC 1344
DB 1702 ATATATGTGCGAGTACCCATTC 1725

RESULT 3
US-09-188-930-67
Sequence 67, Application US/09188930A
Patent No. 6150502
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
APPLICANT: Murlson, James Greg
TITLE OF INVENTION: Compositions Isolated From Skin Cells
FILE REFERENCE: 11000.1011c1
CURRENT APPLICATION NUMBER: US/09/188.930A
CURRENT FILING DATE: 1998-11-09
NUMBER OF SEQ. ID NOS: 348
SOFTWARE: FastSeq for Windows Version 3.0
SEQ. ID NO 67
LENGTH: 1260
TYPE: DNA
ORGANISM: Rat
US-09-188-930-67

Query Match 62.63% Score 841.8; DB 3; Length 1260;
Best Local Similarity 89.43% Pred. No. 2.5e-259;
Matches 900; Conservative 5; Mismatches 102; Indels 0; Gaps 0;

OY 338 GCGCTTTGGATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 397
DB 7 GTGCGTTTGGATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 66
OY 398 CAGATTCACAGAGTGCACACCCAGATCTGATCAATGATGATGATGATGATGATGATGATGAT 457
DB 67 CAGATTCACAGAGTGCACACCCAGATCTGATCAATGATGATGATGATGATGATGATGATGAT 126
OY 458 GCTCTGACACGAGATATGCTTGTGAAAGGCGAGTGTGATGATGATGATGATGATGAT 517
DB 127 GCTCTGACACGAGATATGCTTGTGAAAGGCGAGTGTGATGATGATGATGATGATGATGAT 186
OY 518 GCTATGTTACTGACAGACCTGTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 577
DB 187 GCTATGTTACTGACAGACCTGTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 246
OY 578 ACCCTGTTTACCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 637
DB 247 ACCCTGTTTACCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 306
OY 638 CCGAGAACCCCTGCTGACAACTGCTCAACACTAGGCTCTTTTATCTGCGCTGTG 697
DB 307 CCGAGAACCCCTGCTGACAACTGCTCAACACTAGGCTCTTTTATCTGCGCTGTG 366
OY 698 ACCCAGATATGAACTTGAAGAGATGGCTTCAATGATGATGATGATGATGATGATGATGAT 757
```

Matches 1344: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 ATGCAGAAATAAAGATAGTCTGCTTACATTCCTGCTGCTTCCAAACCC 60
DB 382 ATGCCAGAAATAAAGATAGTCTGCTTACATTCCTGCTGCTTCCAAACCC 441
OY 61 GGAATGACACAGGACAGATGACGAAATGGCTTTGACCTGATCGCAGTACAGCAAGT 120
DB 442 GGAATGACACAGGACAGATGACGAAATGGCTTTGACCTGATCGCAGTACAGCAAGT 501
OY 121 TTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
DB 502 TTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 561
OY 181 AACCAAAATGGCGGATTTATGATTCCTCCGACAAACCTGTGATGAGGGCCCTAC 240
DB 562 AACCAAAATGGCGGATTTATGATTCCTCCGACAAACCTGTGATGAGGGCCCTAC 621
OY 241 TCGAACCCCTACTGACCCCTACTGACGCTCGTACCCGACAGCTGCCACCACTCTCA 300
DB 622 TCGAACCCCTACTGACCCCTACTGACGCTCGTACCCGACAGCTGCCACCACTCTCA 681
OY 301 GCTCCAAATCTCCACAGATCTCCAGGCTCTTATATGCTGCTTGGATACCAATGAT 360
DB 682 GCTCCAAATCTCCACAGATCTCCAGGCTCTTATATGCTGCTTGGATACCAATGAT 741
OY 361 GAAAGCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 742 GAAAGCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 801
OY 421 ACCCAGATCTGATCAATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 802 ACCCAGATCTGATCAATGATGATGATGATGATGATGATGATGATGATGATGAT 861
OY 481 CTTCGAAAGGCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 862 CTTCGAAAGGCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 921
OY 541 TGTGCGAATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 922 TGTGCGAATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 981
OY 601 GATGGAAGTCTTGGCAAGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 982 GATGGAAGTCTTGGCAAGATGATGATGATGATGATGATGATGATGATGATGAT 1041
OY 661 TGGCTCAACACCTGAGCTTTTCACTGCGGCTGTCGACCCGAGATGATGATGAT 720
DB 1042 TGGCTCAACACCTGAGCTTTTCACTGCGGCTGTCGACCCGAGATGATGATGAT 1101
OY 721 GATGCGTTCATTCAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 1102 GATGCGTTCATTCAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1161
OY 781 GAGTGTGTAACAGCCCGGACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
DB 1162 GAGTGTGTAACAGCCCGGACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1221
OY 841 GATGACAACGAGCTGCAAGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 1222 GATGACAACGAGCTGCAAGATGATGATGATGATGATGATGATGATGATGATGAT 1281
OY 901 CTGCGAGAGAGCTGCTCAATTTACAAAGGGGCTTCAATGATGATGATGATGATGAT 960
DB 1282 CTGCGAGAGAGCTGCTCAATTTACAAAGGGGCTTCAATGATGATGATGATGATGAT 1341
OY 961 GAGAGGCTTATCTGAGGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 1342 GAGAGGCTTATCTGAGGATGATGATGATGATGATGATGATGATGATGATGAT 1401
OY 1021 TGCAGAGACGAGCTTACCATCTTACCGGAGATGATGATGATGATGATGATGATGAT 1080
DB 1402 TGCAGAGACGAGCTTACCATCTTACCGGAGATGATGATGATGATGATGATGATGAT 1461

```

```

OY 1081 GTTCCCGGTGACATCTTCCAAATGCAAGCCAGACCGCTACCTGGGCGCTTATCAT 1140
DB 1462 GTTCCCGGTGACATCTTCCAAATGCAAGCCAGACCGCTACCTGGGCGCTTATCAT 1521
OY 1141 TTCCAGATCAAAATCTGGGAATGAGGGCAGAGAAATTTTACATGCGGCAAAAGGGCCCATC 1200
DB 1522 TTCCAGATCAAAATCTGGGAATGAGGGCAGAGAAATTTTACATGCGGCAAAAGGGCCCATC 1581
OY 1201 AGTCCACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
DB 1582 AGTCCACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1641
OY 1261 GAAATGATCACTGTCACACTGTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1320
DB 1642 GAAATGATCACTGTCACACTGTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1701
OY 1321 ATATATGTTGTCAGTACCATTC 1344
DB 1702 ATATATGTTGTCAGTACCATTC 1725

```

RESULT 2

US-09-212-168-2

Sequence 2, Application US/09212168

Patent No. 6303765

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Corley, Neil C.

TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/212,168

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/884,072

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0333 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2550 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: CORN001

CLONE: 45517

US-09-212-168-2

Query Match 100.0%; Score 1344; DB 4; Length 2550;
 Best Local Similarity 100.0%; P-Id No. 0;
 Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 10:57:33 ; Search time 62.0182 Seconds
(without alignments)
6646.008 Million cell updates/sec

Sequence: 1 atgcacgaataaaagat.....atgttcgacgtaccattc 1344

Scoring table: IDENTITY, NUC
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents, NA.*

- 1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
- 4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
- 5: /cgn2_6/ptodata/1/lna/PCUS.COMB.seq:*
- 6: /cgn2_6/ptodata/1/lna/Backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1344	100.0	2550	2	US-08-884-072-2
2	1344	100.0	2550	4	US-09-212-168-2
3	894.8	67.6	1260	3	US-09-188-930-67
4	371.2	27.6	1531	2	US-08-833-963C-1
5	371.2	27.6	2018	3	US-08-980-514-2
6	316.6	23.6	1018	3	US-09-188-930-259
7	241	17.9	2512	4	US-09-248-757-1
8	106.8	7.9	3373	2	US-08-897-443-2
9	89.4	6.7	401	4	US-09-643-597-278
10	80.2	6.0	220	4	US-09-404-879A-340
11	38.2	4.3	2461	1	US-08-282-141-1
12	33.2	4.0	3759	3	US-08-479-7228-3
13	32.4	3.9	5499	3	US-08-479-7228-1
14	52.4	3.9	5502	5	PCT-US95-02251-17
15	51.6	3.8	3753	5	PCT-US95-02251-17
16	51.6	3.8	4314	1	US-08-199-780-2
17	51.6	3.8	4314	2	US-08-316-650-2
18	51.6	3.8	5089	6	5177197-31
19	50.4	3.8	6412	4	US-08-652-877-17
20	50.4	3.8	6412	4	US-08-476-515A-17
21	50.4	3.8	14042	4	US-08-652-877-85
22	50.4	3.8	14044	4	US-08-652-877-89
23	50.4	3.8	14080	4	US-08-652-877-87
24	50.4	3.8	14083	4	US-08-476-515A-83
25	45.6	3.4	3460	2	US-08-652-877-83
26	45.6	3.4	2181	6	US-08-751-305-1
27	44.8	3.3	2181	6	5208144-36

28	43.4	3.2	3546	4	US-08-872-757-3	Sequence 3, Appl
29	41.4	3.1	2492	4	US-09-381-779-1	Sequence 1, Appl
30	39.8	3.0	3136	4	US-09-284-819-8	Sequence 8, Appl
31	38	2.8	6048	4	US-09-634-920-3	Sequence 3, Appl
32	37.8	2.8	1611	4	US-09-249-697A-2	Sequence 2, Appl
33	37.8	2.8	1611	4	US-09-363-316B-2	Sequence 2, Appl
34	37.8	2.8	2365	4	US-09-249-697A-5	Sequence 5, Appl
35	37.8	2.8	2365	4	US-09-249-697A-18	Sequence 18, Appl
36	37.8	2.8	2365	4	US-09-363-316B-5	Sequence 5, Appl
37	37.8	2.8	2365	4	US-09-363-316B-23	Sequence 23, Appl
38	37.2	2.8	2397	6	5258288-2	Patent No. 5258288
39	37	2.8	6344	4	US-08-843-417-1	Sequence 1, Appl
40	35.4	2.6	825	1	US-08-312-870-6	Sequence 6, Appl
41	35.4	2.6	1338	1	US-08-307-444A-8	Sequence 8, Appl
42	35.4	2.6	1338	1	US-08-587-389-8	Sequence 8, Appl
43	35.4	2.6	1368	1	US-08-307-444A-7	Sequence 7, Appl
44	35.4	2.6	1368	1	US-08-587-389-7	Sequence 7, Appl
45	35.4	2.6	1425	1	US-08-307-444A-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-884-072-2
Sequence 2, Application US/08884072
Patent No. 5258234-7

GENERAL INFORMATION:

APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Guebler, Karl J.
TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/884,072
FILING DATE: Herewith
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0333 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:

INFORMATION FOR SEQ. ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 2550 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: CORNNOT01
CLONE: 45517
US-08-884-072-2

Query Match 100.0%; Score 1344; DB 2; Length 2550;
Best Local Similarity 100.0%; Pred. No. 0;

Db	375	GCCTCCAACTACCCCAACGATTTTCAMGGCCTCTTGTCGGCTTTGGGTACAGATGAT	434
QY	361	GAAGCAACCAATGTGTGATGTGACAGATGTGCMAAGATTTCCACCAAGTCAACCC	420
Db	435	GAAAGCAACAGATGTGTGATGTGACAGATGTGCMAAGATTTCCACCAAGTCAACCC	494
QY	421	ACCCACATCTGCATCAATACTGAAGCGGGTACACTGCTCCTGCACCGACGATATTGG	480
Db	495	ACCCACATCTGTATCAACACTGAAGAGGTTACACTGCTCCTGCACCGATGGTACTGG	554
QY	481	CTTCTGGAAGGCGAGTGTAGACATTTGATGATGTCGTTATGTTTACTGCGAGAGTCC	540
Db	555	CTTCTGGAAGGCGAGTGTAGATTTGATGATGTCGTTATGTTTACTGCGAGAGTCC	614
QY	541	TGTGCAATGTTCCTGGATCTATTCCTTTGTACATGCACCCGTGTTTTACCTCATGAG	600
Db	615	TGTGCAATGTTCCTGGATCTATTCCTTTGTACATGCACCCGTGTTTTACCTCATGAG	674
QY	601	GATGGAAGGCTTTGCCAAGATGTGAACGAGTGTGCACGACGACAAACCCGCTGCAAAC	660
Db	675	GATGGAAGGCTTTGCCAAGATGTGAACGAGTGTGCACGACGACAAACCCGCTGCAAAC	734
QY	661	TGCGTCAACACCTACGCGCTCTTTCATCTGCGCGCTGTGACCCAGATTAAGACTGAGAA	720
Db	735	TGCGTCAACACCTATGGCTCTTTCATCTGCGCGCTGTGACCCAGATTAAGACTGAGAA	794
QY	721	GATGGCGTTCAATGGAGATATGACAGATGACAGTCCACTTCTGAGATTCTCTGCCAAT	780
Db	795	GATGGCGTTCAATGGAGATATGACAGATGACAGTCCACTTCTGAGATTCTCTGCCAAT	854
QY	781	GAGTGTGTGMAACAGCCCGGACATACTTCTGCTCCTCCGCTCAGGTACATCTGCTG	840
Db	855	GAGTGTGTGMAACAGCCCGGACATACTTCTGCTCCTCCGCTCAGGTACATCTGCTG	914
QY	841	GATGCAACACCGAAGCTGCCAGACATCAACGAATGTGACAGAGAACACACGTGCAAC	900
Db	915	GATGCAACACCGAAGCTGCCAGAGATTCATGATGTGACACCGAACACACAGTGTACC	974
QY	901	GTGCAGCAGACGTGCTACAAATTTACAAAGGGGCTCAATAGATGATGACCCATCGCGT	960
Db	975	TCACGTGACAGACTGTGTACATCTACAAAGGGGCTCAATAGATGATGATGATGATGAT	1034
QY	961	GAGGAGCCTTATCTGAGATCAGATGATTAACCGCTGTATGTGCTGTGAGAACCTTGCC	1020
Db	1035	GAGGAGCCTTATCTGCTGATGTGTAACACCGCTGTATGTGCTGTGAGAACCTTGCC	1094
QY	1021	TGCAGAGACAGCCCTTTACCATCTTGAACCGGGACATGAGACGTGTGTCAGAGCGTCC	1080
Db	1095	TGCAGAGACAGCCCTTTACCATCTTGAATGGGACATGAGATGTGTGTGTCAGAGCGTCC	1154
QY	1081	GTTCGCCGTGACATCTTCCAAATGCAAGCAGACCGCTACCTGTGGGGCTTATTACAT	1140
Db	1155	GTTCGCCGTGACATCTTCCAGATGCAAGCAGACCGATACCTGTGGGGCTTATTACAT	1214
QY	1141	TTTCCAGATCAATCTGGGAATGAGGGGACGAGAAATTTTATGCGGCAAAAGGGCCCATC	1200
Db	1215	TTTCCAGATCAATCTGGGCAACGAGAGGTGAGAGTTTATTAATGCGGCAAAAGGGCCCATC	1274
QY	1201	AGTGCACACCTGTGATGACACGGCCCATCAAAAGGGCCCGGGAATACACATGGAAGT	1260
Db	1275	AGTGCACACCTGTGATGACACGGCCCATCAAAAGGGCCCGGGAATACACATGGAAGT	1334
QY	1261	GAAATGATCACTGTTCACACCTGTATCACTTCAAGAGGACAGTCCGTGATCCAGTCCGG	1320
Db	1335	GAGATGATCACTGTTCACACCTGTATCACTTCAAGAGGACAGTCCGTGATCCAGTCCGG	1394
QY	1321	ATTATATGTGTGACAGTACCCATTG 1344	
Db	1395	ATTATATGTGTGACAGTATCCGTTG 1418	

AAZ40029	ID	AAZ40029	standard; DNA; 2233 BP.
AAZ40029;	AC	AAZ40029;	
XX	XX	XX	
DT	15-FEB-2000	(first entry)	
XX	XX	XX	
DE	Full length mouse A55 protein coding sequence.		
XX	XX	XX	
KW	A55 protein; mouse; smooth muscle proliferation; tissue generation;		
KW	vascular smooth muscle cell; arteriosclerosis; tissue reparation; myoma;		
KW	vascular endothelial thickening; haemopoietic cell-regulator; cytokine;		
KW	percutaneous transluminal coronary angioplasty; blood coagulation; PTCA;		
KW	actin; inhibin; chemotaxis; thrombosis; cadherin; therapy;		
KW	tumour metastasis inhibitor; ss.		
XX	Mus musculus.		
XX	XX	XX	
PN	WO955864-A1.		
PD	04-NOV-1999.		
XX	XX	XX	
PF	28-APR-1999;	99WO-JP02284.	
XX	28-APR-1998;	98JP-0119731.	
PR	(ONOR) ONO PHARM CO LTD.		
XX	Honjo T, Tashiro K, Nakamura T;		
PI	WPI: 2000-038647/03.		
XX	P-PSDB; AAY54990.		
DR	Novel human polypeptides for treatment of, e.g. arteriosclerosis and		
XX	myoma -		
PT	Example 4; Page 52-55; 87pp; Japanese.		
PS	XX		
XX	XX		
CC	This sequence encodes the mouse A55 protein. The invention relates to the		
CC	human A55 protein. The protein can be used for the treatment of diseases		
CC	due to abnormal proliferation of smooth muscle. The polypeptides can be		
CC	used according their inhibition of the proliferation of vascular smooth		
CC	muscle cells, particularly in treating arteriosclerosis or re-narrowing		
CC	by vascular endothelial thickening after percutaneous transluminal		
CC	coronary angioplasty (PTCA), or myoma, haematopoietic cell-regulatory		
CC	activity, cytokine activity, tissue generation/reparation activity,		
CC	actin/inhibin activity, taxis and chemotaxis activity, blood		
CC	coagulation/thrombotic activity, receptor/ligand activity,		
CC	cadherin/tumour metastasis inhibiting activity; tumour inhibition, and as		
CC	a nutrient.		
XX	Sequence 2233 BP; 554 A; 582 C; 510 G; 586 T; 1 other;		
SO			
Query Match	82.9%; Score 1113.6; DB 21; Length 2233;		
Best Local Similarity	89.3%; Pred. No. 0;		
Matches 1200; Conservative	0; Mismatches 144; Indels 0; Gaps 0		
QY	1 ATGCAGCAATAAAAAGAGTACTGCTACCATCTGGGCTCTGTCCTTCCAGGCCT	60	
DB			
QY	75 ATGCAGAGATTAAAAAGGATCTCTCTTACCACTTGGGCACTGTGGCTTCCACATCCT	134	
DB			
QY	61 GGAATGCACAGGACAGTGCAGCAATGGCTTTGACTGATGCGCCAGTCAAGACAGTGT	120	
DB			
QY	135 GGGATGCACACAGCAGCAGCTGACCAAAAGCGCTTTGACCTGGAGACCGCCAGTCAAGACAGTGT	194	
DB			
QY	121 TTAAATTTGATGTAATGCCCAACCATCCCGCAGGCTCCGAGAGACATGATGTGTGT	180	
DB			
QY	195 CTAAATATTGATGTAATGCCCGGACCATCCCTGAGGCTGTGCTGGGGAATGATGTGTGT	254	
DB			
QY	181 AACCAAAATGGCGGGTATTTTATTCGATTTCCCGGACAAACCTGTGTTATCGAGGGCCCTAC	240	
DB			
QY	255 AACCAAGATGGCGGGTATTTTGTGTGATTCCTCGAACAACCAAGTATATCGAGGGCCCTTAC	314	
DB			

```

QY 421 ACCGAGATTCGATCAATCTGAAAGCGGTACACCTGCTCTGACCGAGCATATG 480
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 ACCGAGATTCGATCAACACTGAAGAGTTACACTGCTCTGACCGAGTACTG 480
QY 481 CTTTGGAGAGCCGAGCTTGTAGACATTCATGATTCGCTATGTTACTGCCAGAGCTC 540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 CTTTGGAGAGGCGAGCTTGTATGATGATGATTCGCTATGTTACTGCCAGAGCTC 540
QY 541 TGTGCAATGTTCTGATCTCTATTCCTGATCAACCCGCTTTTACCCCTCATGAG 600
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 TGTGCAATGTTCTGATCTCTATTCCTGATCAACCCGCTTTTACCCCTCATGAG 600
QY 601 GATGGAAGTCTTCCCAAGATGTGACAGAGTGCACCGAGACCCCTGCGTCAAAAC 660
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 601 GATGGAAGTCTTCCCAAGATGTGACAGAGTGCACCGAGACCCCTGCGTCAAAAC 660
QY 661 TGCCTCAACACCTTACGCTCTTTCATCTCCGCTGTGACCCAGATATGAACTTGAGAA 720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 661 TGTGTCAAACACCTTATGCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGAA 720
QY 721 GATGGCGTTATTCAGTATGATGAGAGTGCACCTCTCTGATGTTCTCTGCCAAT 780
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 721 GATGGCATTCCTGATGATGATGAGAGTGCACCTCTCTGATGTTCTCTGCCAAT 780
QY 781 GAGTGTGTAACCAAGCCCGGACATCTTCTGCTCTGCTCCAGAGTACATCTGCTG 840
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 781 GAGTGTGTAACCAAGCCCGGCTCATCTTCTGCTCTGCTCCAGAGTACATCTGCTG 840
QY 841 GATGACAAACGAGACCTGCAACATCAACGATGTGACAGACAGAACCAACGTCAC 900
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 841 GATGATTAACCGAAGCTGCAAGATATCAATGATGTGACAGACAGAACCAACGTCAC 900
QY 901 CTGACAGCAGAGTCTCAATTTTCAAGGGGCTTCAATGATGATGACCCCATCCGCT 960
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 901 TCACTGAGACTTCTCAATCTTACAGAGGGGCTTCAATGATGATGATGATGATGAT 960
QY 961 GAGGAGCCTTATCTGAGATGAGTAAACCGCTGTATGTCTGCTGAGAACCCCTG 1020
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 961 GAGGAGCCTTATCTGAGATGAGTAAACCGCTGTATGTCTGCTGAGAACCCCTG 1020
QY 1021 TGCAGAGACCAAGCCCTTACATCTTGTACCGGGGACATGAGAGTGTGACAGACCTC 1080
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1021 TGCAGAGACCAAGCCCTTACATCTTGTATGAGGACATGAGATGTGCTGACAGACCTC 1080
QY 1081 GTTCCCGCTGACATCTTCCAAATGCAAGCAGACCCGCTACCTGGGCTATTACAT 1140
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1081 GTTCCCGCTGACATCTTCCAAATGCAAGCAGACCCGCTACCTGGGCTATTACAT 1140
QY 1141 TTTCCAGATCAAAATCTGGGAATGAGGCGAGAAATTTTACATGCGGCAAAAGGCCCATC 1200
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1141 TTTCCAGATCAAAATCTGGCAACGAGGGGTGAGAGTTCTATATGCGGCAAAAGGCCCATC 1200
QY 1201 AGTGGCCACCTGTGTGATGACAGCCCATCAAAAGGGCTCTGGGACATCAGGTGGACTGG 1260
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1201 AGTGGCCACCTGTGTGATGACAGCCCATCAAAAGGGCTCTGGGACATCAGGTGGACTGG 1260
QY 1261 GAATATGATCACTGTCAACACTGTATCACTTCAAGAGGAGAGTCCGTGATCCGACTGCG 1320
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1261 GAGATGATCACTGTCAACACTGTATCACTTCAAGAGGAGAGTCCGTGATCCGACTGCG 1320
QY 1321 ATATATGTGTGCGAGTACCATTC 1344
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1321 ATATATGTGTGCGAGTATCCGCTTC 1344

```

RESULT 14

AAZ39384 standard; DNA: 2233 BP.

ID AAZ39384
 AC AAZ39384;
 XX
 DT 22-FEB-2000 (first entry)

```

XX XX Smooth muscle proliferation modulating protein encoding DNA.
DE Cell proliferation; vascular; smooth muscle cell; arteriosclerosis; PTCA;
XX endothelial thickening; percutaneous transluminal coronary angioplasty;
KW myoma; hematopoietic cell-regulation; cytokine; tissue generation; taxis;
KW actin activity; chemotaxis; blood coagulation; thrombotic; tumor; mouse;
KW metastasis; nutrient; ss.
XX
OS Mus musculus.
XX
PN M0955863-A1.
XX
PD 04-NOV-1999.
XX
PF 28-APR-1999; 99MO-JP02283.
XX
PR 28-APR-1998; 98JP-0119731.
XX
PA (ONOX ) ONO PHARM CO LTD.
XX
PI Honjo T, Tashiro K, Nakamura T;
XX
XX WPI: 2000-038646/03.
DR P-PSDB: AAY56750.
XX
PT Novel mouse polypeptides for treatment, e.g. arteriosclerosis and myoma
PS
PS Claim 5; Page 47-50; 70pp; Japanese.
XX
XX The invention provides mouse polypeptides for treatment of diseases due
CC to abnormal proliferation of smooth muscle. The polypeptides can be
CC produced by standard recombinant methodology. The polypeptides can be
CC used according to their inhibition of the proliferation of vascular
CC smooth muscle cells, particularly in treating arteriosclerosis or re-
CC narrowing by vascular endothelial thickening after percutaneous
CC transluminal coronary angioplasty (PTCA), or myoma, hematopoietic cell-
CC regulatory activity, cytokine activity, tissue generation/reparation
CC activity, actin/inhibitor activity, taxis and chemotaxis activity, blood
CC coagulation/thrombotic activity, receptor/ligand activity, cadherin,
CC tumour metastasis inhibiting activity, tumor inhibition, and as nutrient.
CC The present sequence represents a DNA encoding the protein of the
CC invention which can be used for modulating smooth muscle cell
CC proliferation.
XX
SQ Sequence 2233 BP; 554 A; 582 C; 510 G; 586 T; 1 other:

```

Query Match 82.9%; Score 1113.6; DB 21; Length 2233;
 Best Local Similarity 89.3%; Pred. No. 0;
 Matches 1200; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

```

QY 1 ATGCCAGGAATTAAGATATCTACCTGTTACCATTTGCGCTCTGTCTTCCAACCCCT 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 75 ATGCCAGGAATTAAGATATCTACCTGTTACCATTTGCGCTCTGTCTTCCAACCCCT 134
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GGAATGACAGGACAGTGCACGAAATGAGCTTGAACCTGATGCGCAGTACAGAGAGTGT 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 135 GGAATGACAGGACAGTGCACGAAATGAGCTTGAACCTGATGCGCAGTACAGAGAGTGT 194
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 TTGATATTTGATGATGAGTGCAGAACCATTCGCCAGGCTGCGGAGAGACATGATGTGT 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 195 CTAGATATTTGATGATGAGTGCAGAACCATTCGCCAGGCTGCGGAGAGACATGATGTGT 254
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 AACCAAAATGGCGGATTTTATGATTTCCCGGACAAACCTGTGATGAGAGGCGCTTAC 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 255 AACCAAAATGGCGGATTTTATGATTTCCCGGACAAACCTGTGATGAGAGGCGCTTAC 314
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 TCGAAGCCCTTCTGAGACCCCTTACATGATGCGTACCCGAGAGAGTCCGACACATCTCA 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 315 TCAAAATCCCTTCTTACTACTACTACAGGCCCATATCCAGAGAGGCGGCCACACAGTAC 374
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 GCTCCAAACTATTCAGATCTCCAGGCTCTTATATGCGGCTTGGATACGATGAT 360

```

```

|||||
362 CAGGCTTAGCATGTATGATGATGCGATGTTACTCCAGACACTCTGTCTCGAATGTT 421
QY 553 CCTGGATCTTATTTTGTACTGCAACCCCTGTTTACCCTCAATGAGATGGAAGTCT 612
DB 422 CCTGGATCTTATTTTGTACTGCAACCCCTGTTTACCCTCAATGAGATGGAAGTCT 481
QY 613 TGGCAAGATGTAAGAGAGTGTGCCACCGAGAACCCCTGGTCAACCTGGCTCAACAC 672
DB 482 TGGCAAGATGTAAGAGAGTGTGCCACCGAGAACCCCTGGTCAACCTGGCTCAACAC 541
QY 673 TACGGCTCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGGAAGTGGCTTCAT 732
DB 542 TACGGCTCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGGAAGTGGCTTCAT 601
QY 733 TGCAGTATATGACAGAGTGTGACCTTCTGTAGTTCTCTGCCAATGATGTGTGAA 792
DB 602 TGCAGTATATGACAGAGTGTGACCTTCTGTAGTTCTCTGCCAATGATGTGTGAA 661
Y 793 CAGCCGGACATATCTGTGCTCTGCTGCGCTCAGGCTCATCTGCTGGATGACAACGA 852
DB 662 CAGCCGGACATATCTGTGCTCTGCTGCGCTCAGGCTCATCTGCTGGATGACAACGA 721
QY 853 AGCTGCCAAGACATCAACGAATGTGACACAGAACCAACAGTGCACCTGCAGAGACG 912
DB 722 AGCTGCCAAGACATCAACGAATGTGACACAGAACCAACAGTGCACCTGCAGAGACG 781
QY 913 TGTACTAATTTACAGAGGGGCTTAATGATGATGACACCCCTGCTGTGAGAGCTTAT 972
DB 782 TGTACTAATTTACAGAGGGGCTTAATGATGATGACACCCCTGCTGTGAGAGCTTAT 841
QY 973 CTGAGATCAGTATACCGCTGTATGTGTCTGCTGTGAAACCCCTGGCTGCAGAGACG 1032
DB 842 CTGAGATCAGTATACCGCTGTATGTGTCTGCTGTGAAACCCCTGGCTGCAGAGACG 901
QY 1033 CCTTTACCATCTTGTACCGGGACATGAGCTGTGTGACAGCGCTCCGTTCCCGCTGAC 1092
DB 902 CCTTTACCATCTTGTGTACCGGGACATGAGCTGTGTGACAGCGCTCCGTTCCCGCTGAC 961
QY 1093 ATCTTCCAATGCAAGCGACGACCGCTTACCGTGGGCTTATATCTTTCCAGATCAAA 1152
DB 962 ATCTTCCAATGCAAGCGACGACCGCTTACCGTGGGCTTATATCTTTCCAGATCAAA 1021
QY 1153 TCTGGAAATGAGGGCAGAGATTTTACATGCGGCAAAAGGGGCCCATGATGACACCTG 1212
DB 1022 TCTGGAAATGAGGGCAGAGATTTTACATGCGGCAAAAGGGGCCCATGATGACACCTG 1081
QY 1213 GTGATGACAGCGCCCATCAAAAGGGGCCCGGAAATCCAGCTGACCTGGAATGATCACT 1272
DB 1082 GTGATGACAGCGCCCATCAAAAGGGGCCCGGAAATCCAGCTGACCTGGAATGATCACT 1141
QY 1273 GTCAACACGTCTCATCACTTCAAGAGGACGCTCCGTGATCCGATGCGATATATGTGCG 1332
DB 1142 GTCAACACGTCTCATCACTTCAAGAGGACGCTCCGTGATCCGATGCGATATATGTGCG 1201
QY 1333 CAGTACCATTC 1344
DB 1202 CAGTACCATTC 1213

```

```

KW actin activity; Chemotaxis; blood coagulation; thrombotic; tumor; mouse;
KW metastasis; nutrient; ss.
XX Mus musculus.
XX WO955863-A1.
XX
XX 04-NOV-1999.
XX
XX 28-APR-1999; 99WO-JP02283.
XX
XX 28-APR-1998; 98JP-0119731.
XX
XX (ONOV ) ONO PHARM CO LTD.
XX
XX Honjo T, Tashiro K, Nakamura T;
XX WPI; 2000-038646/03.
XX P-PSDB; AAY56750.
XX
XX Novel mouse polypeptides for treatment, e.g. arteriosclerosis and myoma
XX
XX Claim 4; Page 45-46; 70pp; Japanese.
XX
XX The invention provides mouse polypeptides for treatment of diseases due
XX to abnormal proliferation of smooth muscle. The polypeptides can be
XX produced by standard recombinant methodology. The polypeptides can be
XX used according to their inhibition of the proliferation of vascular
XX smooth muscle cells, particularly in treating arteriosclerosis or re-
XX narrowing by vascular endothelial thickening after percutaneous
XX transluminal coronary angioplasty (PTCA), or myoma, hemipoletic cell-
XX regulatory activity, cytokine activity, tissue generation/repair
XX activity, actin/inhibitor activity, taxis and chemotaxis activity, blood
XX coagulation/thrombotic activity, receptor/ligand activity, cadherin/
XX tumor metastasis inhibiting activity; tumor inhibition, and as nutrient.
XX The present sequence represents the coding sequence of one protein of the
XX invention which can be used for modulating smooth muscle cell
XX proliferation.
XX
XX Sequence 1344 BP; 325 A; 370 C; 330 G; 319 T; 0 other;
XX
XX
XX Query Match 89.9%; Score 113.6; DB 21; Length 1344;
XX Best Local Similarity 82.3%; Pred. NO. 0;
XX Matches 1200; Conservative 0; Mismatches 144; Indels 0; Gaps 0;
XX
XX 1 ATGCCAGGAATTAAGAGATACCTGTTACATCTTGTGCTCTGCTTCCAGACCT 60
XX |||||
XX 1 ATGCCAGGAATTAAGAGATACCTGTTACATCTTGTGCTCTGCTTCCAGATCCT 60
DB |||||
XX 61 GGGAAATGACAGGACAGTGTGACAGATGCTTTGACCTGGATGCCAGTCAAGACGT 120
XX |||||
XX 61 GGGAAATGACAGGACAGTGTGACAGATGCTTTGACCTGGATGCCAGTCAAGACGT 120
DB |||||
XX 121 TTGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
XX |||||
XX 121 CTGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
DB |||||
XX 181 AACCAAAATGAGCGGATATATGATGATGATGATGATGATGATGATGATGATGATGAT 240
XX |||||
XX 181 AACCAAAATGAGCGGATATATGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB |||||
XX 241 TCGAACCCCTGATGACAGCCCTTACTAGTGTGATGATGATGATGATGATGATGATGAT 300
XX |||||
XX 241 TCGAACCCCTGATGACAGCCCTTACTAGTGTGATGATGATGATGATGATGATGATGAT 300
DB |||||
XX 301 GGTCCAAACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
XX |||||
XX 301 GGTCCAAACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
DB |||||
XX 361 GAAAGCAACCAATGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
XX |||||
XX 361 GAAAGCAACCAATGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB |||||

```

RESULT 13
 AA239383
 ID AA239383 standard; DNA; 1344 BP.
 AC AA239383;
 XX
 XX 22-FEB-2000 (first entry)
 DE Smooth muscle proliferation modulating protein coding sequence.
 KW Cell proliferation; vascular; smooth muscle cell; arteriosclerosis; PTCA;
 KW endothelial thickening; percutaneous transluminal coronary angioplasty;
 KW myoma; hematopoietic cell-regulation; cytokine; tissue generation; taxis;

QY	361	GAAGCAACCAATGTGTGAGATGTGAGACAGTGTCAACAGATTTCCACCACTGGCAACCCC	420
Db	742	GAAGCAACCAATGTGTGAGATGTGAGACAGTGTCAACAGATTTCCACCACTGGCAACCCC	801
QY	421	ACCCAGATCTGCATCAATACTGAAGGGGGGTACACCTGCTCTCAGCAGAGATATTGG	480
Db	802	ACCCAGATCTGCATCAATACTGAAGGGGGGTACACCTGCTCTCAGCAGAGATATTGG	861
QY	541	TGTGCGAATGTCTCGATCCATTCTTGTATCTGATATGCAAAACCTGGTTTTACCTCAATAG	600
Db	862	CTTCTGGAAGGCCAGTGTCTTGAACATGTGATTAATGTGGCTATGGTTACTGTGCACAGCTC	921
QY	661	TGCGTCAACACTTACGGCTCTTTTATCTGCGCGCTGTGACCCAGATATGAACCTTGAGAA	720
Db	1042	TGCGTCAACACTTACGGCTCTTTTATCTGCGCGCTGTGACCCAGATATGAACCTTGAGAA	1101
QY	721	GATGCGCTTCAATTCAGATATGAGACGAGTGTGCCACCGAACCCTGCGTGCACAC	780
Db	1102	GATGCGCTTCAATTCAGATATGAGACGAGTGTGCCACCGAACCCTGCGTGCACAC	1161
QY	781	GAGTGTGTGAACACGCCCCGGCACATACTTCTGCTCTGCCCCCTCCAGGCTACATCCTGTG	840
Db	1162	GAGTGTGTGAACACGCCCCGGCACATACTTCTGCTCTGCCCCCTCCAGGCTACATCCTGTG	1221
QY	841	GATGCAACACCGAAGCTGCTCCAAAGATATCAAGAAATGTGAGCAGAGAACACACCTGCAAC	900
Db	1222	GATGCAACACCGAAGCTGCTCCAAAGATATCAAGAAATGTGAGCAGAGAACACACCTGCAAC	1281
QY	901	CTGCGAGCAGAGTGTCTACAAATTTCAAGGGGGGCTTCAAAATGCATTCGACCCATCGCTGT	960
Db	1282	CTGCGAGCAGAGTGTCTACAAATTTCAAGGGGGGCTTCAAAATGCATTCGACCCATCGCTGT	1341
QY	961	GAGGAGCCTTATCTGAGAGATCAGTATPACCGCTGTATGTGCTGTGAGAACCTTGGC	1020
Db	1342	GAGGAGCCTTATCTGAGAGATCAGTATPACCGCTGTATGTGCTGTGAGAACCTTGGC	1401
QY	1021	TGCGAGACACGAGCCCTTTACCATTTTGTACCGGGGACATGAGACGTGTCTCAGAGAGCTCC	1080
Db	1402	TGCGAGACACGAGCCCTTTACCATTTTGTACCGGGGACATGAGACGTGTCTCAGAGAGCTCC	1461
QY	1081	GTTCGCCGCTGACATCTTCCAAATTCGAAGCCACGACCGGCTTCCCTGGGGGCTTATTACAT	1140
Db	1462	GTTCGCCGCTGACATCTTCCAAATTCGAAGCCACGACCGGCTTCCCTGGGGGCTTATTACAT	1521
QY	1141	TTCCAGATCAAAATCTGGGAATGAGGGGACAGAAATTTTACATGCGGCAAAACGGGCCCATC	1200
Db	1522	TTCCAGATCAAAATCTGGGAATGAGGGGACAGAAATTTTACATGCGGCAAAACGGGCCCATC	1581
QY	1201	AGTGCACACCGCTGGATGACACGCCCCCATCAAAAGGGCCCGGGAAATTCACACTGGAATG	1260
Db	1582	AGTGCACACCGCTGGATGACACGCCCCCATCAAAAGGGCCCGGGAAATTCACACTGGAATG	1641
QY	1261	GAATATGATCACTGTCAACACACTGTATCAATCAATTCAGAGGACACTCCGATGATCGACTGGG	1320
Db	1642	GAATATGATCACTGTCAACACACTGTATCAATCAATTCAGAGGACACTCCGATGATCGACTGGG	1701
QY	1321	ATATATGTGTGCGACGATACCATTC	1344
Db	1702	ATATATGTGTGCGACGATACCATTC	1725

XX	AA337670;
XX	
XX	11-SEP-2000 (first entry)
XX	
XX	Human EGF-like homologue (PRO217) encoded by DNA32279 cDNA.
XX	
XX	Inflammatory cell infiltration; immune response; T cell proliferation;
XX	anti-inflammatory; anti-autoimmune; anti-diabetic; spondyloarthropathy;
XX	T cell-mediated disease; spondyloarthropathy; sclerosis; renal disease;
XX	inflammatory myopathy; hemolytic anemia; thrombocytopenia; thyroiditis;
XX	diabetes mellitus; demyelinating polyneuropathy; Guillain-Barre syndrome;
XX	multiple sclerosis; polyneuropathy; hepatitis; cirrhosis; enteropathy;
XX	sclerosing cholangitis; inflammatory bowel disease; Whipple's disease;
XX	skin disease; dermatitis; psoriasis; asthma; allergic rhinitis; tumor;
XX	food hypersensitivity; urticaria; eosinophilic pneumonia; transplant;
XX	idiopathic pulmonary fibrosis; graft rejection; PRO245; human; PRO217;
XX	EGF-like; ss.
XX	
XX	Homo sapiens.
XX	
XX	MO9914241-A2.
XX	
XX	25-MAR-1999.
XX	
XX	17-SEP-1998; 98WO-US19437.
XX	
XX	17-SEP-1997; 97US-0059119.
XX	18-SEP-1997; 97US-0059263.
XX	28-OCT-1997; 97US-0063550.
XX	12-NOV-1997; 97US-0065186.
XX	21-NOV-1997; 97US-0066364.
XX	24-NOV-1997; 97US-0066770.
XX	04-JUN-1998; 98US-0088026.
XX	
XX	(GENE) GENENTECH INC.
XX	
XX	Fong S, Goddard A, Gurney AL, Tumas D, Wood WI;
XX	WPI; 1999-229499/19.
XX	
XX	Composition containing novel polypeptide PRO245, its agonist or
XX	antagonist -
XX	
XX	Example 1; Fig 5A-1-5A-2; 177pp: English.
XX	
XX	This invention describes a novel composition containing (apart from a
XX	carrier or excipient), a novel PRO245 polypeptide (I), its agonist or
XX	antagonist, or their fragments, for modulating: (i) infiltration of
XX	inflammatory cells into tissue; (ii) an immune response; or (iii) T cell
XX	proliferation. The composition increases or decreases any of the effects
XX	(i)-(iii). The products of the invention have anti-inflammatory,
XX	anti-autoimmune and anti-diabetic activity. (I), and its (ant)agonists
XX	and their fragments, are used to treat immune-related diseases,
XX	particularly T cell-mediated diseases. The diseases treated include
XX	systemic lupus erythematosus, rheumatoid arthritis, juvenile chronic
XX	arthritis, spondyloarthropathies, systemic sclerosis (scleroderma),
XX	idiopathic inflammatory myopathies (dermatomyositis, polymyositis),
XX	Sjogren's syndrome, systemic vasculitis, sarcoidosis, autoimmune
XX	hemolytic anemia (immune pancytopenia, paroxysmal nocturnal
XX	hemoglobinuria), autoimmune thrombocytopenia (idiopathic thrombocytopenic
XX	purpura immune-mediated thrombocytopenia), thyroiditis (Grave's disease,
XX	Hashimoto's thyroiditis, juvenile lymphocytic thyroiditis, atrophic
XX	thyroiditis), diabetes mellitus, immune-mediated renal disease
XX	(glomerulonephritis, tubulointerstitial nephritis), multiple sclerosis,
XX	idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome, chronic
XX	inflammatory demyelinating polyneuropathy, infectious hepatitis
XX	(hepatitis A, B, C, D, E and other non-hepatotropic viruses), autoimmune
XX	chronic active hepatitis, primary biliary cirrhosis, granulomatous
XX	hepatitis, and sclerosing cholangitis, inflammatory bowel disease
XX	(ulcerative colitis; Crohn's disease), gluten-sensitive enteropathy, and
XX	Whipple's disease. Autoimmune or immune-mediated skin diseases including
XX	bullous skin diseases, erythema multiforme, contact dermatitis, psoriasis,
XX	


```

OY 361 GAAAGCAACCAATGTGTGATGTGACGAGTGTGCAACAGATTTCCACAGTGCACACCCC 420
DB 546 GAAAGCAACCAATGTGTGATGTGACGAGTGTGCAACAGATTTCCACAGTGCACACCCC 605
OY 421 ACCGAGATCTGCATCAATCTGAAGGCGGGTACACCTGCTCCCTGACACCGAGATATTTGG 480
DB 606 ACCGAGATCTGCATCAATCTGAAGGCGGGTACACCTGCTCCCTGACACCGAGATATTTGG 665
OY 481 CTTGCGAAGGCGGCTGTAGACATTTGAAATGTGCGATGATGTTACTCCAGCAGCTC 540
DB 666 CTTGCGAAGGCGGCTGTAGACATTTGAAATGTGCGATGATGTTACTCCAGCAGCTC 725
OY 541 TGTGCGAATGTCTCTGATCTCTATCTTGTACATGCAACCCGTGTTTACCTTCATGAG 600
DB 726 TGTGCGAATGTCTCTGATCTCTATCTTGTACATGCAACCCGTGTTTACCTTCATGAG 785
OY 601 GATGGAAGGCTTGTCCAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCTGCAAAAC 660
DB 786 GATGGAAGGCTTGTCCAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCTGCAAAAC 845
OY 661 TGGCGTCAACACCTACGCGCTCTTTTCATCTGCGCGCTGTGACCCAGATATGAACTTGAGAA 720
DB 846 TGGCGTCAACACCTACGCGCTCTTTTCATCTGCGCGCTGTGACCCAGATATGAACTTGAGAA 905
OY 721 GATGCGCTTCAATGCAATGATGACGAGTGCAGCTTCTGTGATTTCTCTGCGCAACAT 780
DB 906 GATGCGCTTCAATGCAATGATGACGAGTGCAGCTTCTGTGATTTCTCTGCGCAACAT 965
OY 781 GAGTGTGTAAACAGCGCGGCAATCTTGTGCTCCGCGCTGTGACCGCTACATCTGCTG 840
DB 966 GAGTGTGTAAACAGCGCGGCAATCTTGTGCTCCGCGCTGTGACCGCTACATCTGCTG 1025
OY 841 GATGCAACCGAAGCTGTCCAAAGACATCAACGAATGTGAGACAGAAACACACGTCGAC 900
DB 1026 GATGCAACCGAAGCTGTCCAAAGACATCAACGAATGTGAGACAGAAACACACGTCGAC 1085
OY 901 CTGACGACAGACGTGTCAATTTTCAAGGGGGCTTCAAAATGCAATGCAACCCCATCGCTG 960
DB 1086 CTGACGACAGACGTGTCAATTTTCAAGGGGGCTTCAAAATGCAATGCAACCCCATCGCTG 1145
OY 961 GAGGAGCCTTATGTGAGGATCAGTATACCGCTGTATGTGCTGCTGAGAAACCTGTGC 1020
DB 1146 GAGGAGCCTTATGTGAGGATCAGTATACCGCTGTATGTGCTGCTGAGAAACCTGTGC 1205
OY 1021 TGCAGAGACAGCCCTTTTACCATCTTGTACCGGAGATGACGCTGTCTCAGAGCGCTCC 1080
DB 1206 TGCAGAGACAGCCCTTTTACCATCTTGTACCGGAGATGACGCTGTCTCAGAGCGCTCC 1265
OY 1081 GTTCCCGCTGACATCTTCCAAATGCAAGCCAGACCGCTACCCCTGGGGGCTATTTACATT 1140
DB 1266 GTTCCCGCTGACATCTTCCAAATGCAAGCCAGACCGCTACCCCTGGGGGCTATTTACATT 1325
OY 1141 TTCCAGATCAAAATCTGGGAATGAGGCGAGAAATTTTACATGCGGCAAAACGGGCCCCCATC 1200
DB 1326 TTCCAGATCAAAATCTGGGAATGAGGCGAGAAATTTTACATGCGGCAAAACGGGCCCCCATC 1385
OY 1201 AGTGCACACCTGTGTATGACAGCGCCCATCAAAAGGCCCCCGGAAATCCAGCTGAGCTTG 1260
DB 1386 AGTGCACACCTGTGTATGACAGCGCCCATCAAAAGGCCCCCGGAAATCCAGCTGAGCTTG 1445
OY 1261 GAATGATCACTGTCACACAGCTGTCATCAATTTCAAGAGCAGCTCCGTGATCCGACCTGCG 1320
DB 1446 GAATGATCACTGTCACACAGCTGTCATCAATTTCAAGAGCAGCTCCGTGATCCGACCTGCG 1505
OY 1321 ATATATGTGTGCGAGTACCCATTC 1344
DB 1506 ATATATGTGTGCGAGTACCCATTC 1529

```

RESULT 8
AA05359
ID AAX05359 standard; DNA; 2550 BP.

```

XX AC AAX05359;
XX DT 07-MAY-1999 (first entry)
XX DE Human extracellular matrix protein (ECMP)-1 encoding DNA.
XX DE Extracellular matrix protein; ECMP-1; diagnosis; treatment; cancer;
XX DE Immune disorder; human; ss.
XX OS Homo sapiens.
XX PH Key
XX FT CDS
XX FT Location/Qualifiers
XX FT 382..1728
XX FT /tag="a"
XX FT /product="ECMP-1"
XX PN W09900410-A2.
XX PD 07-JAN-1999.
XX PF 23-JUN-1998; 98MO-US13012.
XX PR 27-JUN-1997; 97US-0884072.
XX PA (INCYTE PHARM INC.
XX PI Bandman O, Corley NC, Guegler KJ;
XX DR WPI; 1999-095674/08.
XX DR P-PSDB; AAW94281.
XX PT New polynucleotide encoding extracellular matrix protein, ECMP-1 -
XX PT useful in the diagnosis, prevention and treatment of immune
XX PT disorders and cancer
XX PS Claim 3; Fig 1A-G; 79pp; English.
XX CC This DNA encodes a human extracellular matrix protein (ECMP)-1. Host
XX CC cells containing a vector comprising the ECMP-1 nucleic acid are used
XX CC for the recombinant production of the protein. ECMP-1 and its
XX CC (ant)agonists, are useful in the diagnosis, prevention, and treatment
XX CC of cancer and immune disorders.
XX SQ Sequence 2550 BP; 623 A; 690 C; 596 G; 641 T; 0 other;

Query Match 100.0%; Score 1344; DB 20; Length 2550;
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;
Matches 1344; Conservative 0; Indels 0; Gaps 0;

OY 1 ATGCCAGGATTAAGGATTAATGATCTGTTACATTTCTGCTCTCTCTTCCAAAGCCT 60
DB 382 ATGCCAGGATTAAGGATTAATGATCTGTTACATTTCTGCTCTCTCTTCCAAAGCCT 441
OY 61 GGGAGATGACAGGACAGTGCAGAAATGCTTTGACCTGATCGGACGAGAGAGTGT 120
DB 442 GGGAGATGACAGGACAGTGCAGAAATGCTTTGACCTGATCGGACGAGAGAGTGT 501
OY 121 TTTGATTTGATGTAATGCGCAACCATCCCGAGGCGCTCCGAGAGAGATGATGTGT 180
DB 502 TTTGATTTGATGTAATGCGCAACCATCCCGAGGCGCTCCGAGAGAGATGATGTGT 561
OY 181 AACCAAAATGCGGCTATTTATGATTCCTCCGCAAAACCTGTGTATGAGGCGCTTC 240
DB 562 AACCAAAATGCGGCTATTTATGATTCCTCCGCAAAACCTGTGTATGAGGCGCTTC 621
OY 241 TCGAACCCCTACTGACCCCTACTACAGTCCGTACCCAGAGCGGCGCCACACACTCA 300
DB 622 TCGAACCCCTACTGACCCCTACTACAGTCCGTACCCAGAGCGGCGCCACACACTCA 661
OY 301 GCTCAACATATCCACAGATCTCCAGGCGCTTATATGCGCTTTGATACAGATGAT 360
DB 682 GCTCAACATATCCACAGATCTCCAGGCGCTTATATGCGCTTTGATACAGATGAT 741

```



```

Db      1009 GATGACACCGAAGCTGCCAAGACATCAACGAATGTGAGCAACAGAACACACGNGAAC 1068
QY      901 CTGACGACAGACGTGCTACAAATTTACAAGGGGGCTTCAAAATGATCGACGCCCTCGCTGT 960
Db      1069 CTGACGACAGACGTGCTACAAATTTACAAGGGGGCTTCAAAATGATCGACGCCCTCGCTGT 1128
QY      961 GAGGAGCCCTTATCTGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Db      1129 GAGGAGCCCTTATCTGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1188
QY      1021 TGCAGAGACAGCCCTTTACCATCTGTACCGGACATGACGATGATGATGATGATGATGATGATGATGAT 1080
Db      1189 TGCAGAGACAGCCCTTTACCATCTGTACCGGACATGACGATGATGATGATGATGATGATGATGATGAT 1248
QY      1081 GTTCCCGCTGACATCTTCCAAATGCAAGCAGCAGCCGCTACCTCGGGGCTATTACATT 1140
Db      1249 GTTCCCGCTGACATCTTCCAAATGCAAGCAGCAGCCGCTACCTCGGGGCTATTACATT 1308
QY      1141 TTCCAGATCAAAATCTGGGAATGAGGGGAGAGAAATTTTACATGCGGCAAAAGGGGCCCATC 1200
Db      1309 TTCCAGATCAAAATCTGGGAATGAGGGGAGAGAAATTTTACATGCGGCAAAAGGGGCCCATC 1368
QY      1201 AGTGCACCCCTGTGTGATGACACGCCCATCAAAAGGGGCCGGAATTCACCTGACCTTG 1260
Db      1369 AGTGCACCCCTGTGTGATGACACGCCCATCAAAAGGGGCCGGAATTCACCTGACCTTG 1428
QY      1261 GAAATATATCACTGATCAACATGTCATCACTTTCAGAGGACGCTCCGTGATCCGATGCGG 1320
Db      1429 GAAATATATCACTGATCAACATGTCATCACTTTCAGAGGACGCTCCGTGATCCGATGCGG 1488
QY      1321 ATATATGTGTGCGAGTACCCATTTC 1344
Db      1489 ATATATGTGTGCGAGTACCCATTTC 1512

```

RESULT 6
AAK07567
ID AAK07567 standard; cDNA; 2362 BP.

XX AC AAK07567;

XX DT 21-JUN-1999 (first entry)

XX DE Homo sapiens fetal kidney clone AK647 secreted protein gene.

XX KW Secreted protein; fetal kidney; ds.

XX DS Homo sapiens.

XX FT Key Location/Qualifiers

XX FT CDS 186..1532 /tag- a /product- "secreted protein"

XX MO9900405-AI.

XX PD 07-JAN-1999.

XX PF 29-JUN-1998; 98MO-US13530.

XX PR 30-JUN-1997; 97US-0885610.

XX PA (GENE) GENETICS INST INC.

XX PI Agostino MJ, Evans C, Jacobs K, Lavalie ER, McCoy JM,

XX PI Merberg D, Racie LA, Treacy M;

XX DR MPI: 1999-095671/08.

XX DR P-PSDB; AAM95709.

XX PT New polynucleotides encoding secreted human proteins - are derived
from foetal kidney or adult retina cDNA libraries, used as, e.g.
potential vaccines

XX PS Disclosure: Pages 51-52; (76pp): English.

XX CC The sequence is that encoding a secreted protein from a human fetal
CC kidney clone AK647. Such a sequence is predicted to have biological
CC activities which would make them suitable for treating, preventing or
CC ameliorating medical conditions in humans and animals, although no
CC supporting data is given. Suggested activities include nutritional
CC activity, cytokine and cell proliferation/differentiation activity,
CC humoral stimulating (e.g. as vaccines) or suppressing activity,
CC haematopoiesis regulating activity, tissue growth activity, hemostatic
CC activin/inhibin activity, chemotactic/chemokinetic activity, anti-inflammatory
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, and tumour
CC inhibition activity. It is also stated to be useful for gene
CC therapy.

SQ Sequence 2362 BP; 580 A; 650-C; 548 G; 584 T; 0 other;

Query Match 100.0%; Score 1344; DB 20; Length 2362;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 ATGCCAGGATTAAGAGATACATCACTGATACATTCCTGCTCTCTTCCAAAGCCT 60
Db      186 ATGCCAGGATTAAGAGATACATCACTGATACATTCCTGCTCTCTTCCAAAGCCT 245
QY      61 GGGAAATGACAGCAGCAGTGACAGAAATGGCTTTGACCTGATGATGATGATGATGATGATGATGATGAT 120
Db      246 GGGAAATGACAGCAGCAGTGACAGAAATGGCTTTGACCTGATGATGATGATGATGATGATGATGATGAT 305
QY      121 TTGATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
Db      306 TTGATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 365
QY      181 AACCAAAATGGCGGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
Db      366 AACCAAAATGGCGGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 425
QY      241 TCGAAGCCCTACTGACACCCCTACTCAGATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 300
Db      426 TCGAAGCCCTACTGACACCCCTACTCAGATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 485
QY      301 GCTCAAAATATCCACAGATGCTCAGGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Db      486 GCTCAAAATATCCACAGATGCTCAGGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 545
QY      361 GAAAGCAACAAATGTGTGATGTGACAGATGTGACAGATGATGATGATGATGATGATGATGATGATGAT 420
Db      546 GAAAGCAACAAATGTGTGATGTGACAGATGTGACAGATGATGATGATGATGATGATGATGATGATGAT 605
QY      421 ACCGAGATGTCATCAATGATGAGAGGGGGGATGACCTGCTGACAGGAGGATGATG 480
Db      606 ACCGAGATGTCATCAATGATGAGAGGGGGGATGACCTGCTGACAGGAGGATGATG 665
QY      481 CTTCGGAAGGACAGTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Db      666 CTTCGGAAGGACAGTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 725
QY      541 TGTGCAATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Db      726 TGTGCAATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 785
QY      601 GATGGAAGTCTTCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db      786 GATGGAAGTCTTCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 845
QY      661 TGCCTCAACACCTAGGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db      846 TGCCTCAACACCTAGGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 905
QY      721 GATGAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780

```

Oy	961	AGAGAGCCTTATCTGAGGATGAGGATTAACCGCTGTATGTCTCGTGAGAAACCTGGC	1020
Dp	1166	GAGAGACCTTATCTGAGGATGAGGATTAACCGCTGTATGTCTCGTGAGAAACCTGGC	1225
Oy	1021	TGCAGAGACCAAGCCCTTATCCATCTTGTACCGGACATGACCTGTGTCTCAGACGCTCC	1080
Dp	1226	TGCAGAGACCAAGCCCTTATCCATCTTGTACCGGACATGACCTGTGTGTCTCAGACGCTCC	1285
Oy	1081	GTTTCCCGCTGACATCTTTCACCAATGCACGACCCGCTACCTGTGGGGCCTATTACATT	1140
Dp	1286	GTTTCCCGCTGACATCTTTCACCAATGCACGACCCGCTACCTGTGGGGCCTATTACATT	1345
Oy	1141	TTTCCAGATCAAAATGTGGGAATGAGGGGACAGAAATTTTACATGCGGGCAAAACGGGCCCATC	1200
Dp	1346	TTTCCAGATCAAAATGTGGGAATGAGGGGACAGAAATTTTACATGCGGGCAAAACGGGCCCATC	1405
Oy	1201	AGTGCACACCCGTGTGATGACACGGCCCATCAAAAGGGCCCCGGGAAATTCACACTGGACTTG	1260
Dp	1406	AGTGCACACCCGTGTGATGACACGGCCCATCAAAAGGGCCCCGGGAAATTCACACTGGACTTG	1465
Oy	1261	GAAATGATCACTGTTCACACACTGTCTCACTCAACTTCAGAGGACAGCTCCGTGATCGACTGGGG	1320
Dp	1466	GAAATGATCACTGTTCACACACTGTCTCACTCAACTTCAGAGGACAGCTCCGTGATCGACTGGGG	1525
Oy	1321	ATTATATGTCTGCGAGTACCCATTTC	1344
Dp	1526	ATTATATGTCTGCGAGTACCCATTTC	1549

RESULT 5
AAZ40027
ID AAZ40027 standard; DNA; 2328 BP.

DT 15-FEB-2000 (first entry)

Full length human A55 protein coding sequence.

KM A55 protein; human; smooth muscle proliferation; tissue generation;
KM vascular smooth muscle cell; arteriosclerosis; tissue reparation; myoma;
KM vascular endothelial thickening; hematopoietic cell-regulator; cytokine
KM percutaneous transluminal coronary angioplasty; blood coagulation; PTCA;
KM actin; inhibin; chemotaxis; thrombosis; cadherin; therapy;
KM tumour metastasis inhibitor; ss.

os Homo sapiens

PN W09955864-A1

04-NOV-1999

PF 28-APR-1999; 99WO-JP02284.

PR 28-APR-1998; 98JP-0119731.

PA (ONOH) ONO PHARM CO LTD.

PI Honjo T, Tashiro K, Nakamura T

DR WPI; 2000-038647/03.

2000

PT myoma -

PS Claim 4, 5; Page 76-80; 87pp; Japanese.

CC This sequence encodes the human A55 protein of the invention. The protein
CC can be used for the treatment of diseases due to abnormal proliferation
CC of smooth muscle. The polypeptides can be used according their inhibition
CC of the proliferation of vascular smooth muscle cells, particularly in
CC treating arteriosclerosis or re-narrowing by vascular endothelial

CC thickening after percutaneous transluminal coronary angioplasty (PTCA),
CC or myoma, haematopoietic cell-regulatory activity, cytokine activity,
CC tissue generation/repairation activity, actin/inhibin activity, taxis
CC and chemotaxis activity, blood coagulation/thrombotic activity,
CC receptor/ligand activity, cadherin/tumour metastasis inhibiting activity;
CC tumour inhibition, and as a nutrient.
CC Note: The open reading frame of this sequence, as well as the mature
CC protein coding sequence are specifically claimed in claim 4.
XX
SQ Sequence 2328 BP, 560 A; 646 C; 541 G; 561 T; 0 other;

Query Match	100.0%	Score 1344	DB 21	Length 2328
Best Local Similarity	100.0%	Pred. NO. 0		
Matches 1344	0	Mismatches	0	Gaps 0

QY	I	ATGCCAGAAATAAAAAGATACTACACTCTTTACACTTCTGCGCTCTGTGCTTCCAGCCCT	60
Db	169	ATGCCAGAAATAAAAAGATACTACTCTTTACACTTCTGCGCTCTGTGCTTCCAGCCCT	228
QY	61	GGAAATGCACAGCGCAGCTGCACGAATGGCTTTCAGCTCGGATCCGCAGCTCAGACAGTGT	120
Db	229	GGGAATGCACAGCGCAGCTGCACGAATGGCTTTCAGCTCGGATCCGCAGCTCAGACAGTGT	288
QY	121	TTAGATATTGTAAGTAATGCCGAACCAATCCCGAGGCGCTCCGAGAGACATGATGTGTCTT	180
Db	289	TTAGATATTGTAAGTAATGCCGAACCAATCCCGAGGCGCTCCGAGAGACATGATGTGTCTT	348
QY	181	AACCAAAATGGCGGCTATTATTAATGCTATCCCGGAAACAAACCCGTGTATACGAGGGCCCTAC	240
Db	349	AACCAAAATGGCGGCTATTATTAATGCTATCCCGGAAACAAACCCGTGTATACGAGGGCCCTAC	408
QY	241	TCGAACCCCTTACTGCAGCCCTCTACTCAGATCCGTACCCAGCAGTCCGCCACACTCTCA	300
Db	409	TCGAACCCCTTACTGCAGCCCTCTACTCAGATCCGTACCCAGCAGTCCGCCACACTCTCA	468
QY	301	GCTCCAAACTATCCACGATCTCCAGGCGCTCTTATATGCGCGTTTGATATCCAGATGTGAT	360
Db	469	GCTCCAAACTATCCACGATCTCCAGGCGCTCTTATATGCGCGTTTGATATCCAGATGTGAT	528
QY	361	GAACACAACCAATGTGTGATGTGGAGAGGTGGACAGATGTGCAACAGATATCCACAGTGCAAACCC	420
Db	529	GAACACAACCAATGTGTGATGTGGAGAGGTGGACAGATGTGCAACAGATATCCACAGTGCAAACCC	588
QY	421	ACCAGATCTGCATCAATACTGTAAGAGCGGGTACACCTCGCTCCCTGCACGACGAGATTTGG	480
Db	589	ACCAGATCTGCATCAATACTGTAAGAGCGGGTACACCTCGCTCCCTGCACGACGAGATTTGG	648
QY	481	CTTCTGGAAGGCCAGTCTTACACTTTGATGATGTCCGTATGTACTGTCACGACGCTC	540
Db	649	CTTCTGGAAGGCCAGTCTTACACTTTGATGATGTCCGTATGTACTGTCACGACGCTC	708
QY	541	TGTGGGATGTTCCGAGATCCGATCTTGTACATGCACAAACCCGGTTTAAACCTCAANTAG	600
Db	709	TGTGGGATGTTCCGAGATCCGATCTTGTACATGCACAAACCCGGTTTAAACCTCAANTAG	768
QY	601	GATGGAAGGTCTTGCCAAAGATGTGAACAGATGTGCCACGAGAACCCCTGCGCTGCAACCC	660
Db	769	GATGGAAGGTCTTGCCAAAGATGTGAACAGATGTGCCACGAGAACCCCTGCGCTGCAACCC	828
QY	661	TGCGTCAACACCTACGGCTCTTTTCAATCTGCGGCTGTGACCAAGATATGAATTAAGAA	720
Db	829	TGCGTCAACACCTACGGCTCTTTTCAATCTGCGGCTGTGACCAAGATATGAATTAAGAA	888
QY	721	GATGGCGTTCAATGTGAGTATGATGAGAGAGTGCACCTCTCTGATTCCTCTCGCAACAT	780
Db	889	GATGGCGTTCAATGTGAGTATGAGAGAGTGCACCTCTCTGATTCCTCTCGCAACAT	948
QY	781	GAGTGTGTGAACCAAGCCCGGACCAATACTTTCGCTCTCGCCCTCCAGGGCTACATCTCTGCTG	840
Db	949	GAGTGTGTGAACCAAGCCCGGACCAATACTTTCGCTCTCGCCCTCCAGGGCTACATCTCTGCTG	1008
QY	841	GATGACAACCGAAGCTGCCAAGACATCAACGAATGTGACACAGGAAACCAACAGCTGCAC	900

QY 1021 TGCAGACAGCCGCTTACCATCTTGTACCGGGAGACGCGTGTGACGACGCTCC 1080
 Db 1231 TGCAGACAGCCGCTTACCATCTTGTACCGGGAGACGCGTGTGACGACGCTCC 1290
 QY 1081 GTTCCCGCTGACATCTTCCAAATGACAGCCAGCCGCTACCTGGGGCTATTACAT 1140
 Db 1291 GTTCCCGCTGACATCTTCCAAATGACAGCCAGCCGCTACCTGGGGCTATTACAT 1350
 QY 1141 TTCCAGATCAATCTGGGAGTGAAGGAGAGAAATTTTACATGCGCAAGCGGCCCATTC 1200
 Db 1351 TTCCAGATCAATCTGGGAGTGAAGGAGAGAAATTTTACATGCGCAAGCGGCCCATTC 1410
 QY 1201 ATGTCCACCTGTGTGATGACAGCGCCCATCAAAAGGGCCCGGAAATCAGCTGGAGCTTG 1260
 Db 1411 ATGTCCACCTGTGTGATGACAGCGCCCATCAAAAGGGCCCGGAAATCAGCTGGAGCTTG 1470
 QY 1261 GAAATGATCACTGTCAACACTGTATCATCACTTCAAGAGGACGCTCGTATCGACTGGCG 1320
 Db 1471 GAAATGATCACTGTCAACACTGTATCATCACTTCAAGAGGACGCTCGTATCGACTGGCG 1530
 QY 1321 ATATATGTGTGCGCATACCATTC 1344
 Db 1531 ATATATGTGTGCGCATACCATTC 1554
 RESULT 4
 AAK94505
 ID AAK94505 standard; cDNA; 2126 BP.
 AC AAK94505;
 XX
 DT 06-NOV-2001 (first entry)
 DE Human full-length cDNA, SEQ ID NO: 3356.
 KM Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
 OS Homo sapiens.
 XX
 PN EP130094-A2.
 PD 05-SEP-2001.
 XX
 PF 07-JUL-2000; 2000EP-0114089.
 XX
 PR 08-JUL-1999; 99JP-0194486.
 PR 11-JAN-2000; 2000JP-0118774.
 PR 02-MAY-2000; 2000JP-0183765.
 XX
 (HELI-) HELIX RES INST.
 PA Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 DR MPI: 2001-524255/58.
 DR P-RSDB; AAM93573.
 XX
 PT 830 Primers useful for synthesizing full length cDNA clones and their
 use in genetic manipulation -
 XX
 PS Claim 8; SEQ ID NO 3356; 1380bp + sequence listing; English.
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesised by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is a full length
 CC human cDNA of the invention.

CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.
 XX

SQ Sequence 2126 BP; 501 A; 608 C; 520 G; 497 T; 0 other;

Query Match
 Best Local Similarity 100.0%; Score 1344; DB 22; Length 2126;
 Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCCAGGATTAATAAGATACATCTTACATCTTGTGCGCTCTGCTTCCAGCCCT 60
 Db 206 ATCCAGGATTAATAAGATACATCTTACATCTTGTGCGCTCTGCTTCCAGCCCT 265
 QY 61 GGAATGACAGGACAGTGCAGCAATGCTTTGACCTGATTCGCAAGCAAGTGT 120
 Db 266 GGAATGACAGGACAGTGCAGCAATGCTTTGACCTGATTCGCAAGCAAGTGT 325
 QY 121 TTGATATATGATGAATCCGAAACCATCCGAGGCTCCGAGAGACATGATGTGT 180
 Db 326 TTGATATATGATGAATCCGAAACCATCCGAGGCTCCGAGAGACATGATGTGT 385
 QY 181 AACCAAAATGGGGGTATTTATGATTCCTCCGCAAAACCTGTATCGAGGGCCCTAC 240
 Db 386 AACCAAAATGGGGGTATTTATGATTCCTCCGCAAAACCTGTATCGAGGGCCCTAC 445
 QY 241 TCGAACCCCTACTGACCCCTTACTAGTCCGATACCGAGAGCTGCCCCACACTCTCA 300
 Db 446 TCGAACCCCTACTGACCCCTTACTAGTCCGATACCGAGAGCTGCCCCACACTCTCA 505
 QY 301 GCTCCAAATATCCCAAGCATCTCAAGGCTCTTATATGCCGTTTGATACCATGAT 360
 Db 506 GCTCCAAATATCCCAAGCATCTCAAGGCTCTTATATGCCGTTTGATACCATGAT 565
 QY 361 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAAGATTTCCACAGTGAACCCC 420
 Db 566 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAAGATTTCCACAGTGAACCCC 625
 QY 421 ACCCAGATCTGATCAATACTGAAGGCGGATACACCTCTCTGACAGCAGATATTGG 480
 Db 626 ACCCAGATCTGATCAATACTGAAGGCGGATACACCTCTCTGACAGCAGATATTGG 685
 QY 481 CTCTGGAAGGCGAGTGTATGACATGTATGATGCTGCTATGCTTACTGCCAGAGCTC 540
 Db 686 CTCTGGAAGGCGAGTGTATGACATGTATGATGCTGCTATGCTTACTGCCAGAGCTC 745
 QY 541 TTGTGCAATGTCTCCGATCTCTATCTTGTATGATCAACCCGTTTACCTCAATGAG 600
 Db 746 TTGTGCAATGTCTCCGATCTCTATCTTGTATGATCAACCCGTTTACCTCAATGAG 805
 QY 601 GATGGAAGTCTTCCCAAGATGTGACAGAGTGTGACACCGAGAACCCCTGCGTCAAAC 660
 Db 806 GATGGAAGTCTTCCCAAGATGTGACAGAGTGTGACACCGAGAACCCCTGCGTCAAAC 865
 QY 661 TGCCTCAACACCTACGCGCTTTCATCTGCGGCTGTGACACCGAGATATCACTGGAG 720
 Db 866 TGCCTCAACACCTACGCGCTTTCATCTGCGGCTGTGACACCGAGATATCACTGGAG 925
 QY 721 GATGCGCTTCAATTCAGTATGATGAGAGAGTGTGACAGTCTGTGAGTCTGCGCAACAT 780
 Db 926 GATGCGCTTCAATTCAGTATGATGAGAGAGTGTGACAGTCTGTGAGTCTGCGCAACAT 985
 QY 781 GATGCTGGAACAGCCCGGACATCTTCTCTGCTGCTGCAAGGCTACATCTGCTG 840
 Db 986 GATGCTGGAACAGCCCGGACATCTTCTCTGCTGCTGCAAGGCTACATCTGCTG 1045
 QY 841 GATGACAACCGAAGCTGCAAGACATCAACGATGTGACACAGAACACAGCTGCAC 900
 Db 1046 GATGACAACCGAAGCTGCAAGACATCAACGATGTGACACAGAACACAGCTGCAC 1105
 QY 901 CTGAGACAGAGCTGTATTAATTTCAAGGGGCTTCAATGATGACCCCATCGCTGT 960
 Db 1106 CTGAGACAGAGCTGTATTAATTTCAAGGGGCTTCAATGATGACCCCATCGCTGT 1165

xx	ABR13627 standard; cDNA: 1720 BP.
xx	
xx	ABK13627;
xx	
xx	23-APR-2002 (first entry)
xx	
xx	Human cDNA encoding extracellular protein-like/EGF-like protein, EGF.
xx	
xx	Human; ss; gene; extracellular protein-like protein; EGF-like;
xx	protein; epidermal growth factor; EGF; ATCC 97285; gene therapy;
xx	vascular smooth muscle cell proliferation; Marfan syndrome;
xx	wound healing; neurological trauma; acquired immunodeficiency syndrome;
xx	AIDS-related dementia; ocular disorders; kidney disorder; liver disorder;
xx	hair follicle growth promotion; burn; ulcer; corneal incision;
xx	corneal inflammation; neoplasm; psoriasis.
xx	
xx	Homo sapiens.
xx	
xx	Key
xx	CDS
xx	Location/Qualifiers
xx	211..1557
xx	/*tag- a
xx	/product= "EGF"
xx	/note= "this CDS minus the stop codon (211-2554) is
xx	specifically claimed in claim 5"
xx	211..285
xx	/*tag- b
xx	286..1554
xx	/*tag- c
xx	/label= Mature_EGF
xx	/note= "this region is specifically claimed in
xx	claim 4"
xx	
xx	US2001051358-A1.
xx	
xx	13-DEC-2001.
xx	
xx	25-MAR-1999; 99US-0275805.
xx	
xx	11-APR-1997; 97US-0839525.
xx	10-APR-1996; 96WO-US05247.
xx	
xx	(OLSE/) OLSEN H. S.
xx	(LIHH/) LI H.
xx	
xx	Olsen HS, Li H;
xx	
xx	WPI; 2002-121417/16.
xx	P-PSDB; AAU75494.
xx	
xx	New nucleic acid encoding human extracellular/epidermal growth factor,
xx	useful for diagnosis and treatment of e.g. Marfan syndrome and wounds,
xx	also related polypeptides
xx	
xx	Claim 4; Fig 1; 22pp; English.
xx	
xx	The invention relates to a novel polynucleotide which is at
xx	least 95% identical with a sequence (ATCC 97285) encoding mature human
xx	extracellular protein-like/epidermal growth factor (EGF)-like protein,
xx	EGF. Also included are the EGF EGF domains, a vector containing
xx	the polynucleotide, a host cell containing the vector, anti-EGF
xx	antibodies and antagonists of EGF. The polynucleotide is used for
xx	recombinant production of EGF, in gene therapy, as hybridisation probes
xx	as antisense antagonists and for chromosome identification. The protein
xx	is used to treat patients who require EGF, to identify specific
xx	antagonists, used to treat conditions that require inhibition of EGF
xx	(e.g. vascular smooth muscle cell proliferation, Marfan syndrome, wound
xx	healing, neurological trauma, acquired immunodeficiency syndrome,
xx	(AIDS)-related dementia, ocular disorders, kidney disorders, liver
xx	disorders, hair follicle growth promotion, burns, ulcers, corneal
xx	incisions, corneal inflammation, neoplasm and psoriasis), to raise
xx	specific antibodies and to characterise receptors. The present
xx	sequence encodes EGF.

Seq	Sequence	1720 BP	396 A	515 C	434 G	372 T	3 other:
	Query Match	100.0%	Score 1344	DB 24	Length 1720		
	Best Local Similarity	100.0%	Mismatches 0				
	Matches 1344	Conservative 0	Mismatches 0	Indels 0	Gaps 0		
QY	1 ATGCAGGAATATAAAGGATATCTCATGTTTACATTCCTGGGCTCTCTGCTCCAAAGCCCT	60					
DB	211 ATGCCAGGAATATAAAGGATATCTCATGTTTACATTCCTGGGCTCTCTGCTCCAAAGCCCT	270					
QY	61 GGAATGCACAGGACACAGTGCACGAATGGCTTTGACCTGATCGCCAGTCGACAGAGTGT	120					
DB	271 GGAATGCACAGGACACAGTGCACGAATGGCTTTGACCTGATCGCCAGTCGACAGAGTGT	330					
QY	121 TTAATATATATGTAATGCCAACCATTCCCGAGGCTCCGAGAGATATATGTGT	180					
DB	331 TTAATATATATGTAATGCCAACCATTCCCGAGGCTCCGAGAGATATATGTGT	390					
QY	181 AACCAAAATGCGGGGTATTTATGCAATTCGCCGACAAACCTGTGTATCGAGGGCCCTTAC	240					
DB	391 AACCAAAATGCGGGGTATTTATGCAATTCGCCGACAAACCTGTGTATCGAGGGCCCTTAC	450					
QY	241 TCGAACCCTTACTGACCCCTTACTCAGGTCCGTACCCAGCAGTGTGCCACCACTCTCA	300					
DB	451 TCGAACCCTTACTGACCCCTTACTCAGGTCCGTACCCAGCAGTGTGCCACCACTCTCA	510					
QY	301 GCTCAAACTATCCACAGATCTCCAGGCTCTTATATGCGCTTTGGATACAGATGGAT	360					
DB	511 GCTCAAACTATCCACAGATCTCCAGGCTCTTATATGCGCTTTGGATACAGATGGAT	570					
QY	361 GAAAGCAACCAATGTGTGATGTGAGAGTGTGCACAGATTCCTCCACCACTGCACACCC	420					
DB	571 GAAAGCAACCAATGTGTGATGTGAGAGTGTGCACAGATTCCTCCACCACTGCACACCC	630					
QY	421 ACCAGATCTGCATCAATACACTGAAAGGGGGGTACACCTGCTCCTGACCGAGGATTTGG	480					
DB	631 ACCAGATCTGCATCAATACACTGAAAGGGGGGTACACCTGCTCCTGACCGAGGATTTGG	690					
QY	481 CTTCTGGAAGGCCAGTGGCTTACAGATTTGATGAATGTGCTATGTGTTACTGCGACAGCTC	540					
DB	691 CTTCTGGAAGGCCAGTGGCTTACAGATTTGATGAATGTGCTATGTGTTACTGCGACAGCTC	750					
QY	541 TGTGCGATATCTCTCGATTCCTATTCCTGTATCATAGCAACCCCTGGTTTAACTCCATAAGAG	600					
DB	751 TGTGCGATATCTCTCGATTCCTATTCCTGTATCATAGCAACCCCTGGTTTAACTCCATAAGAG	810					
QY	601 GATGGAAGGTCTTGCCACAAGATGTACAGAGTGTGCCAACCGGCTGCGTGCACAACC	660					
DB	811 GATGGAAGGTCTTGCCACAAGATGTACAGAGTGTGCCAACCGGCTGCGTGCACAACC	870					
QY	661 TGCCTCAACACCTACGGGCTCTTTCATCTGCGCTGTGTGACCCAGATATGAACCTTGAGGAA	720					
DB	871 TGCCTCAACACCTACGGGCTCTTTCATCTGCGCTGTGTGACCCAGATATGAACCTTGAGGAA	930					
QY	721 GATGGCGCTTACATGAGTATGATGAGAGTGCAGGCTCTCTGATGTTCTCTGGCAACAT	780					
DB	931 GATGGCGCTTACATGAGTATGATGAGAGTGCAGGCTCTCTGATGTTCTCTGGCAACAT	990					
QY	781 GAGTGTGTGAACCAAGCCCGGACATATCTTGTCTCTGCGCCCTCCAGGCTACATCTGTGTG	840					
DB	991 GAGTGTGTGAACCAAGCCCGGACATATCTTGTCTCTGCGCCCTCCAGGCTACATCTGTGTG	1050					
QY	841 GATGACAACCGAAGCTGGCAGACATATCAACGAATGTGACACAGAGAACCAACAGTCGACAC	900					
DB	1051 GATGACAACCGAAGCTGGCAGACATATCAACGAATGTGACACAGAGAACCAACAGTCGACAC	1110					
QY	901 CTGCAAGCAGAGTGTCTACATTTTACAAAGGGGGCTTCAAAATCATGACACCCATCGCTGT	960					
DB	1111 CTGCAAGCAGAGTGTCTACATTTTACAAAGGGGGCTTCAAAATCATGACACCCATCGCTGT	1170					
QY	961 GAGGAGCCTTATCTGAGGATGATGATATACCGCTGTATGTGTCTCTGCTGAGAACCTGTGC	1020					
DB	1171 GAGGAGCCTTATCTGAGGATGATGATATACCGCTGTATGTGTCTCTGCTGAGAACCTGTGC	1230					

/note- "Extracellular/epidermal growth factor-like"

XX MO9846746.1
 XX 22-OCT-1998.
 XX

PF 11-APR-1997; 97WO-US06020.
 XX

PR 11-APR-1997; 97WO-US06020.
 XX

PA (HUMA-) HUMAN GENOME SCI INC.
 XX

PI L1 H, Olsen HS;
 XX

WP1: 1998-568728/48.
 DR P-PSDB; AAW79739.
 XX

PT New Isolated extracellular/epidermal growth factor - used for
 PT regulating vascular smooth muscle cell proliferation, e.g. for
 PT enhancing neurological functions or treating neoplasia and other
 PT disorders.

PS Claim 1a; Fig 1A-D; 62pp; English.
 XX

CC This sequence encodes a novel human extracellular/epidermal growth
 CC factor-like protein, EGF. This protein can be used to regulate
 CC vascular smooth muscle cell proliferation and for restoration or
 CC enhancement of neurological functions diminished as a result of other
 CC damaging pathologies such as AIDS dementia. The protein can also be used
 CC to treat senile dementia, ocular disorders such as corneal inflammation,
 CC for targeting tumor cells, for treating kidney disorders, for liver
 CC regeneration or treating liver dysfunction, for treating wounds including
 CC all cutaneous wounds, corneal wounds, and injuries to the
 CC epithelial-lined hollow organs of the body or resulting from trauma such
 CC as burns, abrasions and cuts as well as from surgical procedures such as
 CC surgical incisions and skin grafting. The polypeptides can also be used
 CC for treating chronic conditions, such as chronic ulcers, diabetic ulcers,
 CC other non-healing (trophic) conditions, to treat Marfan syndrome, to
 CC promote hair follicular development, to stimulate growth and in
 CC vitro and in vivo differentiation of various epidermal and epithelial cells in vitro and in
 CC vitro and to stimulate embryogenesis. Antagonists to EGF can be used to
 CC treat neoplasia such as cancers or tumours, skin disorders such as
 CC psoriasis or corneal inflammation. The products can also be used for
 CC identifying EGF receptors, detection, diagnosis and drug screening.
 CC
 XX Sequence 1720 BP; 396 A; 515 C; 434 G; 372 T; 3 other:
 SQ

Query Match 100.0%; Score 1344; DB 19; Length 1720;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGCCAGATATTAAGATGACTGTTACCTTTGGCTCTCTGCTTCCAAAGCCCT 60
 DB 211 ATGCCAGATATTAAGATGACTGTTACCTTTGGCTCTCTGCTTCCAAAGCCCT 270
 OY 61 GGGAGTGCACAGGACAGTGCAGGAATGGCTTGACCTGCATCCAGTCAAGACAGTGT 120
 DB 271 GGGAGTGCACAGGACAGTGCAGGAATGGCTTGACCTGCATCCAGTCAAGACAGTGT 330
 OY 121 TTGATATTGATGATATGCCAGACCATCCCGAGGCTGCGAGGAGAGACATGATGTGT 180
 DB 331 TTGATATTGATGATATGCCAGACCATCCCGAGGCTGCGAGGAGAGACATGATGTGT 390
 OY 181 AACCAAAATGCGCGGATTTTATGATTCCTCCGCGAACAACCTGTGTATCGAGGCGCTTAC 240
 DB 391 AACCAAAATGCGCGGATTTTATGATTCCTCCGCGAACAACCTGTGTATCGAGGCGCTTAC 450
 OY 241 TCGAACCCCTACGACGCGCTCTACTAGGTCGATACCGAGAGCGGCCACACACTCTCA 300
 DB 451 TCGAACCCCTACGACGCGCTCTACTAGGTCGATACCGAGAGCGGCCACACACTCTCA 510
 OY 301 GCTCAAACTATCCAGATCTCCAGGCTCTTATATGCGCTTGTGATACAGATGAT 360
 DB 511 GCTCAAACTATCCAGATCTCCAGGCTCTTATATGCGCTTGTGATACAGATGAT 570
 OY 361 GAAAGCAACCAATGTGTGATGTGACAGAGTGTGCAACAGATTTCCACAGTGCACACCCC 420
 DB 571 GAAAGCAACCAATGTGTGATGTGACAGAGTGTGCAACAGATTTCCACAGTGCACACCCC 630
 OY 421 ACCCAGATCTGATCAATACGTAAGGCGGCTACCTGCTCCGACACCGAGATATTTGG 480
 DB 631 ACCCAGATCTGATCAATACGTAAGGCGGCTACCTGCTCCGACACCGAGATATTTGG 690
 OY 481 CTTCGGAAGGCGACAGTGTGACATTTGATGATGATGATGATGATGATGATGATGATGAT 540
 DB 691 CTTCGGAAGGCGACAGTGTGACATTTGATGATGATGATGATGATGATGATGATGATGAT 750
 OY 541 TGTGGAATGTTCTTGTGATCTTATCTTGTATGATGATGATGATGATGATGATGATGAT 600
 DB 751 TGTGGAATGTTCTTGTGATCTTATCTTGTATGATGATGATGATGATGATGATGATGAT 810
 OY 601 GATGGAAGGCTTGTGCAAGATGTGAAGAGTGTGCAAGAGTGTGCAAGAGTGTGCAAGAG 660
 DB 811 GATGGAAGGCTTGTGCAAGATGTGAAGAGTGTGCAAGAGTGTGCAAGAGTGTGCAAGAG 870
 OY 661 TCGGTCAACACCTACGCGCTTTTCATCTGCGCGCTGTGACCCAGATATGAATTTAGAGAA 720
 DB 871 TCGGTCAACACCTACGCGCTTTTCATCTGCGCGCTGTGACCCAGATATGAATTTAGAGAA 930
 OY 721 GATGCGCTTATGCAAGTATATGAGAGTGTGCAAGAGTGTGCAAGAGTGTGCAAGAGTGT 780
 DB 931 GATGCGCTTATGCAAGTATATGAGAGTGTGCAAGAGTGTGCAAGAGTGTGCAAGAGTGT 990
 OY 781 GAGTGTGTGTAACAGCGCGGACATCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 DB 991 GAGTGTGTGTAACAGCGCGGACATCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1050
 OY 841 GATGACACCGAAGCTGTGCAAGATGTGAAGAGTGTGCAAGAGTGTGCAAGAGTGTGCAAG 900
 DB 1051 GATGACACCGAAGCTGTGCAAGATGTGAAGAGTGTGCAAGAGTGTGCAAGAGTGTGCAAG 1110
 OY 901 GTTCAGACAGCGGTGTACAAATTTTACAAGGCGGCTTCAAAATGATGACAGCCATCCGCTGT 960
 DB 1111 GTTCAGACAGCGGTGTACAAATTTTACAAGGCGGCTTCAAAATGATGACAGCCATCCGCTGT 1170
 OY 961 GAGGAGCCTTATCTGAGATCACTGATACCGCTGTATGTGTCTGCTGAGAACCTGTGCG 1020
 DB 1171 GAGGAGCCTTATCTGAGATCACTGATACCGCTGTATGTGTCTGCTGAGAACCTGTGCG 1230
 OY 1021 TCGAGAGACAGCGCTTTTACATCTGTACCGGAGATGAGAGTGTGAGAGAGTGTGAGAG 1080
 DB 1231 TCGAGAGACAGCGCTTTTACATCTGTGTACCGGAGATGAGAGTGTGAGAGAGTGTGAGAG 1290
 OY 1081 GTTCGCGGTGACATCTTCCAAATGCAAGCAGCAGCCGCTACCTGGGGCTTATTTACATT 1140
 DB 1291 GTTCGCGGTGACATCTTCCAAATGCAAGCAGCAGCCGCTACCTGGGGCTTATTTACATT 1350
 OY 1141 TTCCAGATCAATCTGGGAATGAGGCGAGAGATTTTACATGCGGCAAAAGGCGCCCTAC 1200
 DB 1351 TTCCAGATCAATCTGGGAATGAGGCGAGAGATTTTACATGCGGCAAAAGGCGCCCTAC 1410
 OY 1201 AGTGCACACCTGTGATGACAGCGCCCATCAAAAGGCGCCGGAATTCAGTGTGACTTGG 1260
 DB 1411 AGTGCACACCTGTGATGACAGCGCCCATCAAAAGGCGCCGGAATTCAGTGTGACTTGG 1470
 OY 1261 GAAATGATCACTGTCAACACTGTGATCAACTTTCAGAGGAGCTCCGTGATCCGACTGCGG 1320
 DB 1471 GAAATGATCACTGTCAACACTGTGATCAACTTTCAGAGGAGCTCCGTGATCCGACTGCGG 1530
 OY 1321 ATATATGTGTGCGAGTACCATTC 1344
 DB 1531 ATATATGTGTGCGAGTACCATTC 1554

RESULT 3
 ABK13627

PT muscle cells expressing cell proliferating growth factor and cellular
 adherence factor
 PS
 XX Example 2; Page 97; 98pp; English.

CC The present invention describes an artificial vascular graft (I)
 CC comprising a synthetic tubular element having a luminal surface coated
 CC with several endothelial cells (ECs) and/or smooth muscle cells (SMCs)
 CC genetically transformed to express at least one cell proliferating growth
 CC factor (CGF) and a cellular adherence factor (CAF). (I) is useful for
 CC replacing or bypassing a portion of a vascular system of an individual,
 CC by implanting (I) into the vascular system of the individual, to form a
 CC fluid communication between the vascular system and the graft, where the
 CC synthetic tubular element of the graft is of an inner cross sectional
 CC area which is substantially equivalent to an inner cross sectional area
 CC of a blood vessel. The present sequence represents the human UP50 cDNA
 CC sequence (also known as a developmental arteries and neural crest
 CC epidermal growth factor (EGF)-like protein (DANCE) gene), which is used
 CC in an example from the present invention.

Sequence 1347 BP; 329 A; 388 C; 327 G; 303 T; 0 other;

Query Match 100.0%; Score 1344; DB 24; Length 1347;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGCCAGAGATTAAGAGATCTACTCTTACCATCTGGCTCTGTCTTCCAGGCT 60
 1 ATGCCAGAGATTAAGAGATCTACTCTTACCATCTGGCTCTGTCTTCCAGGCT 60
 61 GGGAAATGACAGAGCAGAGTGCAGATGGCTTTGACCTGGATGCCAGTACAGACGT 120
 61 GGGAAATGACAGAGCAGAGTGCAGATGGCTTTGACCTGGATGCCAGTACAGACGT 120
 61 GGGAAATGACAGAGCAGAGTGCAGATGGCTTTGACCTGGATGCCAGTACAGACGT 120
 121 TTAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 121 TTAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 121 TTAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 181 AACCAAAATGGGGGATTTATGATGATGATGATGATGATGATGATGATGATGAT 240
 181 AACCAAAATGGGGGATTTATGATGATGATGATGATGATGATGATGATGATGAT 240
 181 AACCAAAATGGGGGATTTATGATGATGATGATGATGATGATGATGATGATGAT 240
 241 TCGAACCCTTACTGACCCCTTACTGACCCCTTACTGACCCCTTACTGACCCCTT 300
 241 TCGAACCCTTACTGACCCCTTACTGACCCCTTACTGACCCCTTACTGACCCCTT 300
 301 GCTCCAACTATCCACGATCTCCACGATCTCCACGATCTCCACGATCTCCACGAT 360
 301 GCTCCAACTATCCACGATCTCCACGATCTCCACGATCTCCACGATCTCCACGAT 360
 361 GAAAGCAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 361 GAAAGCAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 421 ACCGATCTGATCAATCTGAAGGGGGTACACCTGCTCTGACGACGAGATATGG 480
 421 ACCGATCTGATCAATCTGAAGGGGGTACACCTGCTCTGACGACGAGATATGG 480
 421 ACCGATCTGATCAATCTGAAGGGGGTACACCTGCTCTGACGACGAGATATGG 480
 481 CTTTGGAGAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 481 CTTTGGAGAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 481 CTTTGGAGAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 541 TGTGGGAATGTTCTGATCTATCTTGTATGATGATGATGATGATGATGATGATGAT 600
 541 TGTGGGAATGTTCTGATCTATCTTGTATGATGATGATGATGATGATGATGATGAT 600
 541 TGTGGGAATGTTCTGATCTATCTTGTATGATGATGATGATGATGATGATGATGAT 600
 601 GATGGAAGGTCTTGGCAAGATGTGAAGAGTGTGCAACGAGAACCCCTCGTCAAAAC 660
 601 GATGGAAGGTCTTGGCAAGATGTGAAGAGTGTGCAACGAGAACCCCTCGTCAAAAC 660
 661 TGCGCAACACCTACGAGCTCTTCTCATCTGCGCTGTGACCCAGGATATGAACTTGA 720
 661 TGCGCAACACCTACGAGCTCTTCTCATCTGCGCTGTGACCCAGGATATGAACTTGA 720

721 GATGCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 721 GATGCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 781 GATGCTGTGAACACCCCGGACATATCTTGTCTCTGCTCCCTCCAGCTATCTGCTG 840
 781 GATGCTGTGAACACCCCGGACATATCTTGTCTCTGCTCCCTCCAGCTATCTGCTG 840
 841 GATGCAACCCGAGAGTGTCCAAAGATCAAGATGATGATGATGATGATGATGATGATGAT 900
 841 GATGCAACCCGAGAGTGTCCAAAGATCAAGATGATGATGATGATGATGATGATGATGAT 900
 901 CTGACGACAGCTGTCTATCTTACAAAGGGGGCTTCAAAATGATGATGATGATGATGAT 960
 901 CTGACGACAGCTGTCTATCTTACAAAGGGGGCTTCAAAATGATGATGATGATGATGAT 960
 961 GAGGAGCTTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
 961 GAGGAGCTTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
 1021 TGCAGAGACACCCCTTACCAATCTTGTACCCGGACATGAGCGTGTGACAGACGCTCC 1080
 1021 TGCAGAGACACCCCTTACCAATCTTGTACCCGGACATGAGCGTGTGACAGACGCTCC 1080
 1081 GTTCCCGGTGACATCTTCCAAATGCAAGCCAGACCCGCTACCCCTGAGGCTATTACATT 1140
 1081 GTTCCCGGTGACATCTTCCAAATGCAAGCCAGACCCGCTACCCCTGAGGCTATTACATT 1140
 1141 TTCCAGATCAAAATCTGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 1141 TTCCAGATCAAAATCTGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 1201 AGTCCACCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
 1201 AGTCCACCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
 1261 GAAATGATCACTGTCAACACTGTCAATCACTTCAAGAGGAGCTCCGTGATCCGACTCGG 1320
 1261 GAAATGATCACTGTCAACACTGTCAATCACTTCAAGAGGAGCTCCGTGATCCGACTCGG 1320
 1321 ATATATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1344
 1321 ATATATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1344

RESULT 2
 AAV62432
 ID AAV62432 standard; DNA; 1720 BP.
 XX
 AC AAV62432;
 XX
 DT 25-JAN-1999 (first entry)
 XX
 DE Human EGF genomic DNA.
 XX
 KW Extracellular/epidermal growth factor-like protein; EGF; human; liver;
 KW vascular smooth muscle cell proliferation; neurology; pathology; AIDS;
 KW dementia; ocular; disorder; cornea; inflammation; tumour cell; kidney;
 KW wound; epithelium; surgery; skin graft; ulcer; diabetic; trophic;
 KW Marfan syndrome; treatment; hair follicle; embryogenesis; neoplasia;
 KW epidermal cell; cancer; psoriasis; detection; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 211..1557
 FT /tag= a
 FT sig_peptide 211..285
 FT /tag= b
 FT mat_peptide 286..1557
 FT /tag= c
 FT /product= "EGF"

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 08:40:41 ; Search time 299.483 Seconds
(Without alignments)
10106.379 Million cell updates/sec

Database # 1
Sequence: 1 atgcccagagataaaagat.....atgtctcagctaccatc 1344

Scoring table: IDENTITY-NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database # 1
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1344	100.0	1347	24	ABL52528
2	1344	100.0	1720	19	AAV62432
3	1344	100.0	1720	24	ABK13627
4	1344	100.0	2126	22	AAK94505
5	1344	100.0	2328	21	AAZ40027
6	1344	100.0	2362	20	AAZ07567
7	1344	100.0	2362	21	AAZ39892
8	1344	100.0	2550	20	AAZ05359
9	1342.4	99.9	2509	20	AAZ37670

10	1342.4	99.9	2609	22	AAZ46128
11	1342.4	99.9	2609	22	AAZ46128
12	1207.2	89.8	1228	18	AAZ89380
13	1113.6	82.9	1344	21	AAZ39383
14	1113.6	82.9	2233	21	AAZ39384
15	1113.6	82.9	2233	21	AAZ40029
16	1098.6	81.7	1383	21	AAZ39386
17	1098.6	81.7	2429	21	AAZ39387
18	1098.6	81.7	2429	21	AAZ40031
19	1056.2	78.6	1269	21	AAZ39385
20	1056.2	78.6	1269	21	AAZ39388
21	843.8	62.8	1254	21	AAZ61672
22	843.8	62.8	1254	22	AAZ61672
23	841.8	62.6	1260	24	ABL34757
24	776.4	57.8	1778	20	AAZ37673
25	721.4	53.7	853	22	AAZ31596
26	721.4	53.7	853	22	AAZ3159
27	721.4	53.7	1538	22	AAZ94129
28	636.4	47.4	680	21	AAZ44245
29	390	29.0	670	22	AAZ92195
30	390	29.0	670	22	AAZ93463
31	374.4	27.9	377	20	AAZ37686
32	371.2	27.6	1480	21	AAZ14608
33	371.2	27.6	1525	18	AAZ8974
34	371.2	27.6	1531	20	AAZ60351
35	371.2	27.6	1707	22	AAZ13771
36	371.2	27.6	1874	21	AAZ58583
37	371.2	27.6	1875	22	AAZ21402
38	371.2	27.6	1875	24	AAZ18229
39	371.2	27.6	1994	21	AAZ39800
40	371.2	27.6	2018	21	AAZ14603
41	347.2	25.8	1513	21	AAZ14616
42	337.6	25.1	1358	20	AAZ37683
43	326	24.3	354	20	AAZ86404
44	317	23.6	359	20	AAZ61786
45	316.6	23.6	1018	21	AAZ61786

ALIGNMENTS

RESULT 1
ABL52528
ID ABL52528 standard; cDNA; 1347 BP.
AC ABL52528;
DT 19-JUL-2002 (first entry)
DE Human UP50 cDNA sequence SEQ ID NO:1.
XX
XX Human; UP50; artificial vascular graft; cellular adherence factor;
KW cell proliferating growth factor; DANCE; epidermal growth factor; EGF;
KW developmental arteries and neural crest EGF-like protein; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX W0200207646-23.
XX
PD 31-JAN-2002.
XX
XX 20-JUL-2001; 2001MO-1100670.
XX
XX 20-JUL-2000; 2000US-0620227.
XX
XX (MGVS-) MGVS LTD.
XX
XX Flugelman MW, Preiss M, Gluzman Z, Koren B, Weisz A, Cohen T;
XX WPI; 2002-361632/39.
XX
XX Artificial vascular graft for therapy of vascular diseases, has
PT synthetic tubular element having a surface coated with endothelial,


```
Db 632 GATTATTCATTCCTCCGACAAACCTGTGTATGAGGGCCCTACTGAAACCCCTACTC 691
QY 423 GACCCCTACTCAGSTCCGTACCCAGAGCTGCCACACTCTCAGCTCCAAACTATCC 482
Db 692 GACCCCTACTCAGSTCCGTACCCAGAGCTGCCACACTCTCAGCTCCAAACTATCC 751
QY 483 CACGATCTCCAGGCTCTTATATGCGCTTGTGATACCATGATGATGAAGCAACCAATG 542
Db 752 CACGATCTCCAGGCTCTTATATGCGCTTGTGATACCATGATGATGAAGCAACCAATG 811
QY 543 TGTGATGTGAGAGAGTGTGCACAGATCCACAGTGCACACCCAGCCAGATCTGCAT 602
Db 812 TGTGATGTGAGAGAGTGTGC -AMAGATTCACACAGTGCAMCCACCCAGAGTCTGCAT 870
QY 603 CAATACGAAGGGGGGTACACCTGCTGACCGAGGATATGCGCTTGTGAAGGCCA 662
Db 871 CAATACGAAGGGGGGTACACCTGCTGACCGAGGATATGCGCTTGTGAAGGCCA 930
QY 663 GTGCTTAGACATTGATGAATGTGCTATGCTTACTGCCAGCAGCTCTGTGGAATGTCC 722
Db 931 GTGCTTAGACATTGATGAATGTGCTATGCTTACTGCCAGCAGCTCTGTGGAATGTCC 990
QY 723 TGGATCTATCTTGT 738
Db 991 TGGATCTATCTTGT 1006
```

Search completed: July 3, 2003, 17:36:06
Job time : 3436.81 secs

/clone="IMAGE:5229519"
 /clone.lib="NIH_MGC_120"
 /lab.host="DH10B"
 /note="Organ: pooled pancreas and spleen. Vector:
 pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of spleen and pancreas from 28 yo
 male. Library is oligo-dT primed and directionally cloned
 (EcoRV site is destroyed upon cloning). Average insert
 size 1.5 kb, insert size range 1-2.5 kb. Library is
 normalized and enriched for full-length clones and was
 constructed by C. Gruber (Invitrogen). Research Genetics
 tracking code 025. Note: this is a NIH_MGC Library."

BASE COUNT 184 a 240 c 195 g 175 t

ORIGIN

Query Match 30.1%; Score 700.4; DB 13; Length 794;
 Best Local Similarity 98.4%; Pred. No. 4.8e-185;
 Matches 781; Conservative 0; Mismatches 6; Indels 7; Gaps 7;

28 CACGACTCGCTCGCCCTCTGGAATAAACACCCGAGCCGAGGCGCCAGAGAG 87
 1 CACGACTCGCTCGG -CCCTCTGGAATAAACACCCGAGGCGCCAGAGAGAG 59
 88 CCGAGCTGCCGAGCTCTCCGGGGGTCCGCCCGC -GAGCTTCTTCTCGCTTGCAT 146
 60 CCGAGCTGCCGAGCTCTCCGGGGGTCCGCCCGCCTGAGCTTCTTCTCGCTTGCAT 119
 147 CTCTCTCTCGGCGGTCTGGACATGCCAGGAATAAAGATGACTGTTACCATCT 206
 120 CTCTCTCTCGGCGGTCTGGACATGCCAGGAATAAAGATGACTGTTACCATCT 179
 207 GGCCTCTCTCTTCCAAAGCCCTGGGAATGCACAGCAGTGCAGATGGCTTGCAT 266
 180 GGCCTCTCTCTTCCAAAGCCCTGGGAATGCACAGCAGTGCAGATGGCTTGCAT 239
 267 GGATGCCGAGTACAGACAGTGTATGATGATGATGATGATGATGATGATGATGAT 326
 240 GGATGCCGAGTACAGACAGTGTATGATGATGATGATGATGATGATGATGATGAT 299
 327 CCGAGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 386
 300 CCGAGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 359
 387 CCCTGTGTATCGAGGGCCCTACTCGAAGCCCTACTCGAAGCCCTACTCGAAGCC 446
 360 CCCTGTGTATCGAGGGCCCTACTCGAAGCCCTACTCGAAGCCCTACTCGAAGCC 419
 447 AACAGTGGCCCACTCTCAGCTCCAAACTAT -CCGAGATCTCCAGGCTCTTATAT 505
 420 AACAGTGGCCCACTCTCAGCTCCAAACTAT -CCGAGATCTCCAGGCTCTTATAT 479
 506 GCGGCTTGGATACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 565
 480 GCGGCTTGGATACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 539
 566 CAGATTCACACAGTGC -AACCCACCCAGATCTCATCTCATCTCATCTCATCTCAT 624
 540 CAGATTCACACAGTGC -AACCCACCCAGATCTCATCTCATCTCATCTCATCTCAT 599
 625 TGCCTCTGACAGCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 684
 600 TGCCTCTGACAGCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 659
 685 CG -CTATGTTACTGACAGCAG -CTCTGTGCGAATGTT -CCTGATCCTATCTTGTACA 741
 660 CGCCTATGTTACTGACAGCAGCTCTGTGCGAATGTT -CCTGATCCTATCTTGTACA 719
 742 TGCACACCTGTTTACTCTCAATGAGATGAGAGGTCTTGCCAGATGTAGACAGTGT 801
 720 TGCACACCTGTTTACTCTCAATGAGATGAGAGGTCTTGCCAGATGTAGACAGTGT 779
 802 GCCACCGAGAACCC 815
 ||||||||| ||

Db 780 GCCACCGAGAGACC 793

RESULT 15
 AL545549 1006 bp mRNA linear EST 16-FEB-2001
 LOCUS AL545549 LIT.NFL006.PL2 Homo sapiens cDNA clone CS001015YE20 5
 DEFINITION prime, mRNA sequence.
 ACCESSION AL545549
 VERSION AL545549.1 GI:12878031
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1006)
 AUTHORS Li, W.B., Gruber, C., Jeesee, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
 source
 1..1006
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS001015YE20"
 /clone.lib="LIT.NFL006.PL2"
 /issue_type="placenta"
 /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6
 vector. Library was normalized. Library was constructed by
 Life Technologies. Contact: Feng Liang Life Technologies,
 a division of Invitrogen 9800 Medical Center Drive
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : fliang@lifetech.com url :
 http://fulllength.invitrogen.com"

BASE COUNT 215 a 300 c 258 g 228 t 5 others

ORIGIN

Query Match 30.1%; Score 700.4; DB 9; Length 1006;
 Best Local Similarity 98.2%; Pred. No. 5.2e-185;
 Matches 723; Conservative 5; Mismatches 6; Indels 2; Gaps 2;

3 CCGGCGCTCTCCCGGTCTCTCTCCAGCATCTGCTGGCCCTCTGGAATAAACACC 62
 273 CCGGCGCTCTCCCGGTCTCTCTCCAGCATCTGCTGGCCCTCTGGAATAAACACC 332
 63 GCGAGCCCGGAGGCGCCAGAGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 122
 333 GCGAGCCCGGAGGCGCCAGAGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 392
 123 CGAGCTTCTTCTCGCTTGCATCTCTCTCGGCGCTTGTGACATGCCAGATATA 182
 393 CGAGCTTCTTCTCGCTTGCATCTCTCTCGGCGCTTGTGACATGCCAGATATA 452
 183 AAGGATACCTACTGTACCATTTGCTCTCTCTTCCAGCCCTGGAGATGACAGGC 242
 453 AAGGATACCTACTGTACCATTTGCTCTCTCTTCCAGCCCTGGAGATGACAGGC 512
 243 AAGGATACCTACTGTACCATTTGCTCTCTCTTCCAGCCCTGGAGATGACAGGC 302
 513 AAGGATACCTACTGTACCATTTGCTCTCTCTTCCAGCCCTGGAGATGACAGGC 572
 303 ATGCCGAAACATCCCGAGGCTCCGAGAGACATGATGATGATGATGATGATGATG 362
 573 ATGCCGAAACATCCCGAGGCTCCGAGAGACATGATGATGATGATGATGATGATG 631
 363 GATTTATGCAATTCCTCCGAGAACCTGTGTATGAGGCGCTTACTGACACCTACT 422
 ||||||||| ||

LOCUS BM974454 735 bp mRNA linear EST 21-MAR-2002
DEFINITION UI-CF-EC1-aca-j-15-0-UI-s1 UI-CF-EC1 Homo sapiens cDNA clone
ACCESSION UI-CF-EC1-aca-j-15-0-UI 3', mRNA sequence.
VERSION BM974454
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 735)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 FORWARD
PolyA=yes.
FEATURES
source Location/Qualifiers
1..735
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-CF-EC1-aca-j-15-0-UI"
/clone_1lb="UI-CF-EC1"
/tissue_type="Lung"
/dev_stage="Adult and Fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-CF-EC1 is a normalized cDNA library containing the
following tissue(s): Normal lung from adult and from fetal
day 64, day 87, week 19 and week 42. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
AAGCGCTTAC.
TAG_LIB=UI-CF-EC1
TAG_TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
and 380-383
TAG_SEQ=AAGTGGTTAC"
BASE COUNT 236 a 139 c 163 g 197 t
ORIGIN
Query Match 30.8%; Score 717.4; DB 14; Length 735;
Best Local Similarity 99.9%; Pred. No. 8e-190;
Matches 718; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1610 TAGGCAATTCCTGCTGACGTTTCCCGAAGAGTACAGCCCGACTTCGACTCTACCT 1669
|||||
DB 734 TAGGCAATTCCTGCTGACGTTTCCCGAAGAGTACAGCCCGACTTCGACTCTACCT 675
QY 1670 GTACTATTGCAGACCTGTACCCCTGCAGAGCTTGCACCCCGAGTTCCATGATAGATT 1729
|||||
DB 674 GTACTATTGCAGACCTGTACCCCTGCAGAGCTTGCACCCCGAGTTCCATGATAGATT 615

QY 1730 ATCAAAAGATTATATCTTCCTCCCGTGAATGAGAGATTGTTGTAATTTTCAAGGCTT 1789
|||||
DB 614 ATCAAAAGATTATATCTTCCTCCCGTGAATGAGAGATTGTTGTAATTTTCAAGGCTT 555
QY 1790 CAGTTTATTTCCACTATTTTCAAGAAATAGATTAGTTGCGGGGCTCGAGTCTATG 1849
|||||
DB 554 CAGTTTATTTCCACTATTTTCAAGAAATAGATTAGTTGCGGGGCTCGAGTCTATG 495
QY 1850 TTCAAGACTGTGAACAGCTTGTGTCACCTTCACCTTCACCTTCCTGCTACCTG 1909
|||||
DB 494 TTCAAGACTGTGAACAGCTTGTGTCACCTTCACCTTCACCTTCCTGCTACCTG 435
QY 1910 TGTACTGCTTTTCAAGAACCCGGAGCTGCGGGGAACCTGGAGTAGTACTTGGCT 1969
|||||
DB 434 TGTACTGCTTTTCAAGAACCCGGAGCTGCGGGGAACCTGGAGTAGTACTTGGCT 375
QY 1970 TTTTGGCTACACAGAGAGCTATGTAAACAAACACAGAGATGCAAGGCTTTTGA 2029
|||||
DB 374 TTTTGGCTACACAGAGAGCTATGTAAACAAACACAGAGATGCAAGGCTTTTGA 315
QY 2030 GAATGTGTTTCAAAACCATGCTGATATTTTCAACCATTAAGAGTTTCAGTTGCTT 2089
|||||
DB 314 GAATGTGTTTCAAAACCATGCTGATATTTTCAACCATTAAGAGTTTCAGTTGCTT 255
QY 2090 AATTTGTATACGGTTTATTTGCTGCTTGTATTTGATTTTGAATTTTAAATATATGTC 2149
|||||
DB 254 AATTTGTATACGGTTTATTTGCTGCTTGTATTTGATTTTGAATTTTAAATATATGTC 195
QY 2150 GTAGATTTCTTCAAGAGCCTTCACAGACATGCTATGTCCTGCCCAACCCAGTGC 2209
|||||
DB 194 GTAGATTTCTTCAAGAGCCTTCACAGACATGCTATGTCCTGCCCAACCCAGTGC 135
QY 2210 TCCTCTCAATTTTACGCCAGTGTGTTTCTTTGAGAGACCCCTTAACTTGTCTTTAGAA 2269
|||||
DB 134 TCCTCTCAATTTTACGCCAGTGTGTTTCTTTGAGAGACCCCTTAACTTGTCTTTAGAA 75
QY 2270 TTTTACCCCAATTTGATTTGGAATGAGAGCTTCCCAACGATTAATATTGTAAGA 2328
DB 74 TTTTACCCCAATTTGATTTGGAATGAGAGCTTCCCAACGATTAATATTGTAAGA 16
RESULT 14
BI837271 794 bp mRNA linear EST 04-OCT-2001
LOCUS BI837271
DEFINITION 603090563P1 NIH_MGC-120 Homo sapiens cDNA clone IMAGE:5229519 5',
mRNA sequence.
ACCESSION BI837271
VERSION BI837271.1 GI:15948809
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 794)
NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M11577 row: n column: 16
High quality sequence stop: 781.
FEATURES
source Location/Qualifiers
1..794
/organism="Homo sapiens"
/db_xref="taxon:9606"

```

Db      547 TCAAAAAGATATATCATGTCCTCCCGATAGAAATGTTGCTGATTTTCAAGCCCTTC 488
QY      1791 ACTTATTTTCCACTATTTTCAAAAAGAAATAGATTAGTTGGGGGCTGTAGTATGT 1850
Db      487 ACTTATTTTCCACTATTTTCAAAAAGAAATAGATTAGTTGGGGGCTGTAGTATGT 428
QY      1851 TCAAAAGAGTGAAGCTGCTGTCCTCTTCTTCCACTCTTCCACTCTCTCTCTCTCTCT 1910
Db      427 TCAAAAGAGTGAAGCTGCTGTCCTCTTCTTCCACTCTTCCACTCTCTCTCTCTCT 368
QY      1911 GTTACTGCTTTTGGCAAGACCCGGGAGCTGGCGGGAAACCTTGGAGTAGTATGTTGCTT 1970
Db      367 GTTACTGCTTTTGGCAAGACCCGGGAGCTGGCGGGAAACCTTGGAGTAGTATGTTGCTT 308
QY      1971 TTTGGCTACAGAGAGGCTATGTATTAACAACACACAGCAGAGATCGAAGGTTTAAAG 2030
Db      307 TTTGGCTACAGAGAGGCTATGTATTAACAACACACAGCAGAGATCGAAGGTTTAAAG 248
QY      2031 AATGTTTCAAAAACCATGCTGCTGATTTTCAACCATAAAGAAAGTTTCACTGTCCTTA 2090
Db      247 AATGTTTCAAAAACCATGCTGCTGATTTTCAACCATAAAGAAAGTTTCACTGTCCTTA 188
QY      2091 AATTTGTATACGCTTAAATTTCTGCTTGTTCATTTTGAATTTTAAAAAATATGTCG 2150
Db      187 AATTTGTATACGCTTAAATTTCTGCTTGTTCATTTTGAATTTTAAAAAATATGTCG 128
QY      2151 TAGAATTCCTTGAAGGCTTACAGACACATGCTATGTCGTCTGCCAACCAGTCT 2210
Db      127 TAGAATTCCTTGAAGGCTTACAGACACATGCTATGTCGTCTGCCAACCAGTCT 68
QY      2211 CCTCTCATTTTGAAGGCTTACAGACACATGCTATGTCGTCTGCCAACCAGTCT 2270
Db      67 CCTCTCATTTTGAAGGCTTACAGACACATGCTATGTCGTCTGCCAACCAGTCT 8
QY      2271 TTTTACC 2277
Db      7 TTTTACC 1

```

RESULT 12
B1523975
LOCUS 721 bp mRNA linear EST 29-AUG-2001
DEFINITION 603052129F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5201767 5',
B1523975
ACCESSION B1523975
VERSION B1523975.1 GI:15348767
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 721)
AUTHORS NIH-MGC http://mgi.ncl.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgrabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M11505 row: j column: 08
High quality sequence stop: 716.
Location/Qualifiers
1. 721
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5201767"
/clone_11b="NIH_MGC_122"

```

/lab host="DH108"
/note="Organ: pooled lung and spleen; Vector: pCMV-Sport6;
Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 026. Note:
this is a NIH-MGC Library."
BASE COUNT      155 a      175 g      158 t
ORIGIN
Query Match      30.9% Score 719.4; DB 13; Length 721;
Best Local Similarity 99.9% Pred. No. 2.2e-190;
Matches 720; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      3 CCGGGGCTCTCCCGGTGCTCTCCAGACTGCTGCGCCCTTGGAAATAAACACC 62
Db      1 CCGGGGCTCTCCCGGTGCTCTCCAGACTGCTGCGCCCTTGGAAATAAACACC 60
QY      63 GCGAGCCCGGAGGCGCCAGAGAGGCGAGCTGCCGAGCTCTCCGGGGTCCGCCG 122
Db      61 GCGAGCCCGGAGGCGCCAGAGAGGCGAGCTGCCGAGCTCTCCGGGGTCCGCCG 120
QY      123 CGAGCTTTCTTCGCGCTTGGCATCTCTCTCGCGGCTTGGCATGCGGAAATAA 182
Db      121 CGAGCTTTCTTCGCGCTTGGCATCTCTCTCGCGGCTTGGCATGCGGAAATAA 180
QY      183 AAGGATACATCATGTTACATTTCTGCTCTCTTCCAGCCCTGGGAATGACAGCC 242
Db      181 AAGGATACATCATGTTACATTTCTGCTCTCTTCCAGCCCTGGGAATGACAGCC 240
QY      243 ACACTGCGACAATGCTTTGACCTGATCCGACAGTCTTTAGATTTAGTATGA 302
Db      241 ACACTGCGACAATGCTTTGACCTGATCCGACAGTCTTTAGATTTAGTATGA 300
QY      303 ATGCCGAACCATCCCGAGGCTCCGAGAGACATGATGTGTAAACAAATGCGCG 362
Db      301 ATGCCGAACCATCCCGAGGCTCCGAGAGACATGATGTGTAAACAAATGCGCG 360
QY      363 GATATTATGATTTCCCGGACAAACCTGTATGAGAGGCGCTACTCGAACCCCTACTC 422
Db      361 GATATTATGATTTCCCGGACAAACCTGTATGAGAGGCGCTACTCGAACCCCTACTC 420
QY      423 GACCCCTACTCAGAGTCCGACAGAGAGTCCGACACATCTCAGCTCCAAACTATCC 482
Db      421 GACCCCTACTCAGAGTCCGACAGAGAGTCCGACACATCTCAGCTCCAAACTATCC 480
QY      483 CAGCATCTCAGAGGCTTTATATGCTCGCTTGGATACAGATGATGAAGAACCAATG 542
Db      481 CAGCATCTCAGAGGCTTTATATGCTCGCTTGGATACAGATGATGAAGAACCAATG 540
QY      543 TGTGATGTGAGAGTGTGCAACAGATTTCCACAGTGCACACCCACCAGATCTGAT 602
Db      541 TGTGATGTGAGAGTGTGCAACAGATTTCCACAGTGCACACCCACCAGATCTGAT 600
QY      603 CAATACGAAGGCGGTACACCTCTCTGACGAGAGATTTGGCTTGTGAAGGCCA 662
Db      601 CAATACGAAGGCGGTACACCTCTCTGACGAGAGATTTGGCTTGTGAAGGCCA 660
QY      663 GTGCTTACATGATTAATGCTGATGTTAGTCTGACAGAGCTCTGTGCAATGTTCC 722
Db      661 GTGCTTACATGATTAATGCTGATGTTAGTCTGACAGAGCTCTGTGCAATGTTCC 720
QY      723 T 723
Db      721 T 721

```

RESULT 13
BM974454/c

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:

http://image.lnl.gov

Plate: L1M13580 row: h column: 08

High quality sequence stop: 668.

Location/Qualifiers

1. 835

/organism="Homo sapiens"

/db.xref="taxon:9606"

/clone="FM6:6187063"

/clone_lib="Lupski_sym pathetic_trunk"

/sex="male"

/tissue_type="sympathetic trunk"

/dev_stage="adult, 16 yr"

/lab_host="DH10B"

/note="Vector: PCW-SPORE (Life Technologies); Site:1;

NotI; Site:2; SalI; cDNA made by oligo-dT priming.

Directionally cloned using the following adaptors:

5'-TGCACCCAGCGCTCG-3' and

5'-GACATGTCATGACGAGCGCGCCCTT(15)-3'. Size selected >

1 kb for average insert length 1.9 kb. This is a primary

library non-amplified. Library constructed by Life

Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor

College of Medicine); available through Life

Technologies."

BASE COUNT 181 a 266 c 206 g 182 t

ORIGIN

Query Match 33.7%; Score 785; DB 14; Length 835;

Best Local Similarity 99.0%; Pred. No. 9e-209;

Matches 811; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

3 CCCGGCGCTCTCCCGCTGCTCTCTCCAGACTGCTCGGCGCCCTCGAATAAACGCC 62

5 CCCGGCGCTCTCCCGCTGCTCTCTCCAGACTGCTCGGCGCCCTCGAATAAACGCC 64

63 GCGAGCCCCGAGGCGCCAGAGAGGCGCGAGCTGCGGAGCTCTCCGGGGTCCCGCCG 122

65 GCGAGCCCCGAGGCGCCAGAGAGGCGCGAGCTGCGGAGCTCTCCGGGGTCCCGCCG 124

123 CGAGCTTCTTCTGCGCTTCCGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 182

125 CGAGCTTCTTCTGCGCTTCCGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 184

183 AAGGATCTCACTGTATTCATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 242

185 AAGGATCTCACTGTATTCATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 244

243 ACACTGCAAGATGGCTTGAAGCTGATGCGCAGTCAAGAGAGTGTATGATTTGATGA 302

245 ACACTGCAAGATGGCTTGAAGCTGATGCGCAGTCAAGAGAGTGTATGATTTGATGA 304

303 ATGCGGAACCATCCCGGAGGCTGCGGAGAGACATGATGTGTTAACCATAATGGCGG 362

305 ATGCGGAACCATCCCGGAGGCTGCGGAGAGACATGATGTGTTAACCATAATGGCGG 364

363 GATTTATGCAATCCCGGAGGCTGCGGAGAGACATGATGTGTTAACCATAATGGCGG 422

365 GATTTATGCAATCCCGGAGGCTGCGGAGAGACATGATGTGTTAACCATAATGGCGG 424

423 GACCCCTACTCAGTCCGATCCAGAGCTGCGGAGAGACATGATGTGTTAACCATAATGG 482

425 GACCCCTACTCAGTCCGATCCAGAGCTGCGGAGAGACATGATGTGTTAACCATAATGG 484

483 CAGGATCTCAGGCTCTTATATGCGCTTTGGATACAGATGATGAAGAACCAAGT 542

485 CAGGATCTCAGGCTCTTATATGCGCTTTGGATACAGATGATGAAGAACCAAGT 544

543 TGTGATGTGAGACAGTGTGCAACGATTTCCACAGTCAACCCCAACCAATGCAAT 602

545 TGTGATGTGAGACAGTGTGCAACGATTTCCACAGTCAACCCCAACCAATGCAAT 604

603 CAATTAAGAGGCGGTACACCTGCTCTGACAGCAGAGATGATGCTTCTGGAAGGCCA 662

|||||

DB 605 CAATTAAGAGGCGGTACACCTGCTCTGACAGCAGAGATGATGCTTCTGGAAGGCCA 664

663 GTGCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 722

665 GTGCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 724

723 TGGATCTTATTTCTTGTACATGCAACCTGCTTTTACCTCAATG-AGATGGAAGTCTT 781

725 TGGATCTTATTTCTTGTACATGCAACCTGCTTTTACCTCAATGAGATGGAAGTCTT 784

782 GCCAAGATGGAACGAGTGTG-TCGACCGGAGAACCCCTG 818

785 GCCAAGATGGAACGAGTGTG-TCGACCGGAGAACCCCTG 823

DB

RESULT 11

AL570066/c

LOCUS

DEFINITION AL570066 LTL_NFL006.Pl2 Homo sapiens cDNA clone CS0D1006YB07 3

ACCESSION AL570066

VERSION AL570066.1 GI:12926024

KEYWORDS

SOURCE

ORGANISM

human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 727)

Li, W.-B., Gruber, C., Jessup, J., and Polyes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

Bp 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr.

Location/Qualifiers

1. 727

/organism="Homo sapiens"

/db.xref="taxon:9606"

/clone="CS0D1006YB07"

/clone_lib="LTL_NFL006.Pl2"

/tissue_type="Placenta"

/note="Vector: PCWSPORT 6; Site:1: NotI; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the PCWSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact: Feng Liang Life Technologies,

Rockville, Maryland 20850, USA Fax: (41) 301 610 8371

Email: fliang@life.com URL: http://fulllength.invitrogen.com"

BASE COUNT 226 a 147 c 166 g 188 t

ORIGIN

Query Match 31.1%; Score 723.8; DB 9; Length 727;

Best Local Similarity 99.7%; Pred. No. 1.3e-191;

Matches 725; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1551 GGCACCAAGGAG 1610

727 GGCACCAAGGAG 668

1611 AGGCAATTCCTGCTGAGAGTTTCCCGAAGAGTCAAGCCGACTTCTGACTCTCACTG 1670

667 AGGCAATTCCTGCTGAGAGTTTCCCGAAGAGTCAAGCCGACTTCTGACTCTCACTG 608

1671 TACATATTGACAGACTGTACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1730

607 TACATATTGACAGACTGTACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 548

1731 TCATAAAGTATTATCATGTCTCCCTGTATGAGAGATTTGTTGATTTTCAAGGCTTC 1790

Db	721	TTTCGTCCTCCCTCCAGGCTACATCCCTGCTGATGATCAACCGGAAGCTCCAAAGAC	780
QY	1032	CATCAAGCAATGTGAGCGACAGCAACACAGGTGCA--CCTGACGACAGCGTCAATT	1099
Db	781	ATCAACCGAATGTAGCGACAGCAACACAGCTGCACCTCTGACAGACGTCCTCAATT	840
QY	1091	T-----ACAAGGGGGCTTCAATGCATGCACCCCATCCGCTGTGAGGAG	1134
Db	841	TTACACAGGGGGCCCTTCAAATGCATGCAGAACCCCATCCCGCTGTGAG	887

RESULT	9
B1764020	
LOCUS	880 bp mRNA linear EST-25-SEP-2001
DEFINITION	603043222.F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5183693 5' ,
ACCESSION	B1764020
VERSION	B1764020
KEYWORDS	B1764020.1 GI:15755598
SOURCE	EST.
ORGANISM	human.
	Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 880)
NIH-MGC <http://mgc.ncl.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

FEATURES
source
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
plate: LLNL1458 row: 1 column: 06
High quality sequence stop: 846.
Location/Qualifiers
1..880

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1IMAGE:5183693"
/clone_1ID="N1H.MGC_116"
/lb_host="DH10B"
/notes="organ: pooled colon, kidney, stomach; Vector:
pCMW-SpORF6; Site_1: NOLr; Site_2: EcoRV (destroyed); RNA
source anonymous; pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a N1H.MGC Library."

```

Query Match	34.0%	Score 790.6	DB 13	Length 880
Best Local Similarity	98.7%	Pred. No. 2.5e-210		
Matches 860	Conservative 0	Mismatches 4	Indels 7	Gaps 6

QY	3	CCGGGCGCTCTCCCCCGTCTCTTCACAGACTCGGTGGGGCCCTGTGAATTTAAACACC	62
Db	11	CCGGGCGCTCTCCCCGTCTCTTCACAGACTCGGTGGGGCCCTGTGAATTTAAACACC	70
QY	63	GCGAGCCCGGAGGGCCGAGAGAGGCCACGTGCCGAGCTCTTCGGGGGTCCCGCCG	122
Db	71	GCGAGCCCGGAGGGCCGAGAGAGGCCGAGTGGCCGAGCTCTTCGGGGGTCCCGCCG	130
QY	123	CGAGCTTCTCTTCGCGCTTGCATCTCTCTCGGCGCTCTTGACATGCCAGATATAA	182

Dp	131	CGAGTTTCTTCTGCGCTTCGATCTCTCCCTCGGCGCTTTGGAAATCGAAGGATATA	190
Oy	183	AAGGATACTCACTGTTACCAATTCGGCTCTGTCTTCTCCAAAGCCTTGGAAATGCACAGCG	242
Dp	191	AAGGATACACTGTTACCATTCTGGCTCTGTCTTCTCCAAAGCCTTGGAAATGCACAGCG	250
Oy	243	ACAGTGCACGAATGGCTTTGACCTGGATTCGCCAGTCAGGACAGTGTTTTGATATTGATGA	302
Dp	251	ACAGTGCACGAATGGCTTTGACCTGGATTCGCCAGTCAGGACAGTGTTTTGATATTGATGA	310
Oy	303	ATGCGAACAACATCCCGGAGAGGCTGCGCGAGAGACATGATGTGTGTTAAACCAAAATGGCG	362
Dp	311	ATGCGAACAACATCCCGGAGAGGCTGCGCGAGAGACATGATGTGTGTTAAACCAAAATGGCG	370
Oy	363	GTAATTATGATTCCTCCCGGACAAACCTGTGTATGAGAGGGCCCTACTCGAACCCCTACTC	422
Dp	371	GTAATTATGATTCCTCCCGGACAAACCTGTGTATGAGAGGGCCCTACTCGAACCCCTACTC	430
Oy	423	GACCCCTTACTCAGGTCCGTCACCCAGCAGCTGCCACACATCTCAAGCTCCAAACTATCC	482
Dp	431	GACCCCTTACTCAGGTCCGTCACCCAGCAGCTGCCACACATCTCAAGCTCCAAACTATCC	490
Oy	483	CACGATTCACAGGCGCTTTATATGGCGCTTTGGATACAGATGGATGAAGAACAACCAATG	542
Dp	491	CACGATTCACAGGCGCTTTATATGGCGCTTTGGATACAGATGGATGAAGAACAACCAATG	550
Oy	543	TGTGGATGTGACGAGTGTGCACACAGATTCGCCACAGTCACAAACCCACCCAGATCTGCAT	602
Dp	551	TGTGGATGTGACGAGTGTGCACACAGATTCGCCACAGTCACAAACCCACCCAGATCTGCAT	610
Oy	603	CAATACGTAAGGCGGGGTACCTGCTCTGTGCACACGACGATATTGGCTTCTGGAAGCCA	662
Dp	611	CAATACGTAAGGCGGGGTACCTGCTCTGTGCACACGACGATATTGGCTTCTGGAAGGCCA	670
Oy	663	GTGCTTACACATGATGATGATGCGGTATGATGATGATGATGATGATGATGATGATGATGATG	721
Dp	671	GTGCTTACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	730
Oy	722	CTGATCTCTATTCCTGTGA - CATGCAACCCCTGGTATTACCTCTCAATGAGAGATGGA - GGT	778
Dp	731	CTGATCTCTATTCCTGTGAATTTGCAACCTGGTATTACCTCTCAATGAGAGATGGAAGGT	790
Oy	779	CTTGGCAAGA - TGTGAACGAGTGTGCAACCGAGAACCCCTGGTGTGAACATCTGCTCAAC	837
Dp	791	CTTGGCAAGATGTGAACGAGTGTGCAACCGAGAACCCCTGGTGTGAACATCTGCTGTCAC	848
Oy	838	ACCTACGAGCTTTATCTTCCTGCGCGCTGTGAC	868
Dp	849	ACCTACGAGCTTTATCTTCCTGCGCGCTGTGAC	879

RESULT 10			
LOCUS	B0718885		
DEFINITION	B0718885	835 bp	mRNA
	ACENSCOURT_8241132	lupsk1-sympathetic-trunk	linear
	IMAGE:6187063_5'	EST 16-JUL-2000	
		Homo sapiens cDNA clone	

VERSION	KEYWORDS
BQ7188885.1	GI:21857782
SOURCE	ORGANISM
human.	Homo sapiens
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	TITLE
1 (bases 1 to 835)	NIH-MGc http://mgc.ncbi.nlm.nih.gov/ .
AUTHORS	JOURNAL
National Institutes of Health, Mammalian Gene Collection (MGC) unpublished (1999)	Contact: Robert Strausberg, Ph.D.
COMMENT	

Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation

QY 1247 CCGTCCCGCTGACATCTTCCAAATGCAAGCCAGACCCGCTACCTGGGGCTATTAC 1306
| | | | |
Db 301 CCGTCCCGCTGACATCTTCCAAATGCAAGCCAGACCCGCTACCTGGGGCTATTAC 360
| | | | |
QY 1307 TTTTCCAGATCAAAATCTGGGAATGAGGGCGAGAAATTTTACATCCGGCAAGGGCCCCA 1366
| | | | |
Db 361 TTTTCCAGATCAAAATCTGGGAATGAGGGCGAGAAATTTTACATCCGGCAAGGGCCCCA 420
| | | | |
QY 1367 TCAGTGGCCACCCCTGGGATGACACGGCCCATCAAGGGCCCCGGGAATTCAGCTGACT 1426
| | | | |
Db 421 TCAGTGGCCACCCCTGGGATGACACGGCCCATCAAGGGCCCCGGGAATTCAGCTGACT 480
| | | | |
QY 1427 TGGAAATGATCACTGTCAACACTGTCTATCAACTTCAGAGAGAGCTCCGTATCCGACTGC 1486
| | | | |
Db 481 TGGAAATGATCACTGTCAACACTGTCTATCAACTTCAGAGAGAGCTCCGTATCCGACTGC 540
| | | | |
QY 1487 GGATATATGTGTGCGAGTACCCATCTTGAGCCCTGGGGCTGGAGCCCTCCGAGCTGCTCT 1546
| | | | |
Db 541 GGATATATGTGTATGCGAGTACCCATCTTGAGCCCTGGGGCTGGAGCCCTCCGAGCTGCTCT 600
| | | | |
QY 1547 CATTGGCAACAGGGAGAGAGAGAGAGAAATTAACAGAGCAATGAGAGAGAGAGAGAG 1605
| | | | |
Db 601 CATTGGCAACAGGGAGAGAGAGAGAGAAATTAACAGAGAGAGAGAGAGAGAGAGAGAG 660
| | | | |
QY 1606 ACCTTGGCACTTCTGCTGAGAGCTTTCCCGAGAGAGTACAGCCGACTTCTGACTCTC 1665
| | | | |
Db 661 ACCTTGGCACTTCTGCTGAGAGCTTTCCCGAGAGAGTACAGCCGACTTCTGACTCTC 720
| | | | |
QY 1666 ACCTTGGCACTTCTGCTGAGAGCTTTCCCGAGAGAGTACAGCCGACTTCTGACTCTC 1724
| | | | |
Db 721 ACCTTGGCACTTCTGCTGAGAGCTTTCCCGAGAGAGTACAGCCGACTTCTGACTCTC 780
| | | | |
QY 1725 CAGTATCAAAAAGATTTATCATTTGCTCCCTGATGAGAGATTTGTTGATTAATTTGA 1784
| | | | |
Db 781 CAGTATCAAAAAGATTTATCATTTGCTCCCTGATGAGAGATTTGTTGATTAATTTGA 839
| | | | |
QY 1785 GCCTTCACTTATTTCCACTATTTTCAAGAAATAGATTAGTTTGGCGG 1835
| | | | |
Db 840 GCCTTCCG -CAATTCCCTATTTCCAGAGACATGATGAGAGAGG 888
| | | | |

RESULT 8
B0716569
LOCUS 904 bp mRNA linear EST 16-JUL-2002
DEFINITION AGENCOURT_8215248 Lupskl_sympathetic_trunk Homo sapiens cDNA clone
IMAGE:6187340 5', mRNA sequence.
B0716569
ACCESSION B0716569.1 GI:21855466
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 904)
AUTHORS NIH-MBC <http://nigc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupskl
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LHAM1581 row: c column: 21
High quality sequence stop: 646.
Location/Qualifiers
1..904

FEATURES
SOURCE
1..904
/organism="Homo sapiens"
/db_xref="taxon:9606"

/clone="IMAGE:6187340"
/clone_lib="Lupskl_sympathetic_trunk"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6 (Life Technologies); Site 1:
NotI; Site 2: SalI; cDNA made by oligo-dT priming.
5'-TCGACCCAGCCGTCG-3' and
5'-GACTGATCTTGAATGCGAGCGGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupskl, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."
BASE COUNT 219 a 264 c 219 g 202 t
ORIGIN
Query Match 34.0%; Score 792.2; DB 14; Length 904;
Best Local Similarity 97.1%; Pred. No. 8.0e-211;
Matches 861; Conservative 0; Mismatches 18; Indels 8; Gaps 5;
256 GCCTTGGACCTGGATGCCAGTCAAGACAGTGTATGATTTGATGATGCCGGAACATC 315
| | | | |
Db 1 GCCTTGGACCTGGATGCCAGTCAAGACAGTGTATGATTTGATGATGCCGGAACATC 60
| | | | |
QY 316 CCGGAGCGCTGCCGAGAGAGATGATGTGTTTACCAAAATGGCGGATTTATGACTT 375
| | | | |
Db 61 CCGGAGCGCTGCCGAGAGAGATGATGTGTTTACCAAAATGGCGGATTTATGACTT 120
| | | | |
QY 376 CCGGAGCAAAACCTGTATGAGAGGCGCTTACCGAACCCTTCTGAGCCCTTACTCA 435
| | | | |
Db 121 CCGGAGCAAAACCTGTATGAGAGGCGCTTACCGAACCCTTCTGAGCCCTTACTCA 180
| | | | |
QY 436 GGTCCGTACCAAGACAGCTGCCGCCACCACTGACCTTCAAACTATCCAGATCTCCAG 495
| | | | |
Db 181 GGTCCGTACCAAGACAGCTGCCGCCACCACTGACCTTCAAACTATCCAGATCTCCAG 240
| | | | |
QY 496 CCTCTTATTTGGCGCTTTGGATGCCAGATGATGAAAGCAACCAATGTGGATGTGAC 555
| | | | |
Db 241 CCTCTTATTTGGCGCTTTGGATGCCAGATGATGAAAGCAACCAATGTGGATGTGAC 300
| | | | |
QY 556 GAGTGTCAACAGATTTCCACAGTCAACCCACAGATGATGATGATGATGATGATGATG 615
| | | | |
Db 301 GAGTGTCAACAGATTTCCACAGTCAACCCACAGATGATGATGATGATGATGATGATG 360
| | | | |
QY 616 GGTACACCTGCTCTGACACCGAGATATTTGGCTTGTGAAGGCCAGTGTAGACATT 675
| | | | |
Db 361 GGTACACCTGCTCTGACACCGAGATATTTGGCTTGTGAAGGCCAGTGTAGACATT 420
| | | | |
QY 676 GATGAATGCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 735
| | | | |
Db 421 GATGAATGCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
| | | | |
QY 736 TGTACATGCAACCTGCTTTTACCTCAATGAGATGAGAGGCTTCCCAAGATGTGAAC 795
| | | | |
Db 481 TGTACATGCAACCTGCTTTTACCTCAATGAGATGAGAGGCTTCCCAAGATGTGAAC 540
| | | | |
QY 796 GAGTGTGCCACCGAGAACCCCTGCGTGAACACCTGCGTCAACACCTTATGATC 855
| | | | |
Db 541 GAGTGTGCCACCGAGAACCCCTGCGTGAACACCTGCGTCAACACCTTATGATC 600
| | | | |
QY 856 TGGCCGCTGAGACCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 915
| | | | |
Db 601 TGGCCGCTGAGACCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
| | | | |
QY 916 GAGTGCAGCTTCTGAGTCTCTGCAACATGATGATGATGATGATGATGATGATGATGAT 975
| | | | |
Db 661 GAGTGCAGCTTCTGAGTCTCTGCAACATGATGATGATGATGATGATGATGATGATGAT 720
| | | | |
QY 976 TTTGCTCTCTGCTCCAGAGCTACAT -CCTGCTGATGATGATGATGATGATGATGATGAT 1031
| | | | |

source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH-MGC Library."

BASE COUNT 191 a 277 c 216 g 197 t 1 others

Query Match 34.7% Score 807.6; DB 13; Length 882;
Best Local Similarity 98.4%; Pred. No. 4.2e-215;
Matches 857; Conservative 0; Mismatches 10; Indels 4; Gaps 4;

1 GACCCGCGCTCCCGCTGCTCCAGACGCTGCGGCGCTCGGATTAACAC 60
11 GATCCGCGCTCCCGCTGCTCCAGACGCTGCGGCGCTCGGATTAACAC 70
61 CCGCGAGCCCGAGGCGCCAGAGAGGCGGAGCTCCGCGGCGGCGCGCC 120
71 CCGCGAGCCCGAGGCGCCAGAGAGGCGGAGCTCCGCGGCGGCGCGCC 130
121 CCGGAGCTTCTCTCCCTTCCGATCTCTCTCCGCGCTTGGACATGCCAGATA 180
131 CCGGAGCTTCTCTCCCTTCCGATCTCTCTCCGCGCTTGGACATGCCAGATA 190
181 AAAAGTACTACTGTTACATCTGGCTCTGCTTCCAGGCCCGGAGGACAG 240
191 AAAAGTACTACTGTTACATCTGGCTCTGCTTCCAGGCCCGGAGGACAG 250
241 GCACAGTGCAGATGCGCTTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 300
251 GCACAGTGCAGATGCGCTTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 310
301 GAATGCCGAAACATCCCGAGGCTCCGAGAGACATGATGTTTAAACAAATGCG 360
311 GAATGCCGAAACATCCCGAGGCTCCGAGAGACATGATGTTTAAACAAATGCG 370
361 GGGTATTTATGATCCCGAGGCTCCGAGAGACATGATGTTTAAACAAATGCG 420
371 GGGTATTTATGATCCCGAGGCTCCGAGAGACATGATGTTTAAACAAATGCG 430
421 TCGACCCCTTACTGAGTCCGTACCCAGACAGCTCCGAGAGACATGATGTTTAAACAAATGCG 480
431 TCGACCCCTTACTGAGTCCGTACCCAGACAGCTCCGAGAGACATGATGTTTAAACAAATGCG 490
481 CCCAGATCTCCAGGCTCTTATATGCGCTTGGATCCAGATGATGAAGCAACAA 540
491 CCCAGATCTCCAGGCTCTTATATGCGCTTGGATCCAGATGATGAAGCAACAA 550
541 TGTGTGATGTGAGCGATGTCAGCAAGATCCAGCGTGCAGCAACCCAGATCTGC 600
551 TGTGTGATGTGAGCGATGTCAGCAAGATCCAGCGTGCAGCAACCCAGATCTGC 610
601 ATCAATATGAAAGCGGCTACCTGCTCCGACCCAGATGATGCTTGAAGGC 660
611 ATCAATATGAAAGCGGCTACCTGCTCCGACCCAGATGATGCTTGAAGGC 670
661 CAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
671 CAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 730
721 CCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 778
731 CCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 790
779 CTTGCAAGATGTAAGC-AGTGTGCAACCGAGAGAGAGAGAGAGAGAGAGAGAG 837
791 CTTGCAAGATGTAAGC-AGTGTGCAACCGAGAGAGAGAGAGAGAGAGAGAGAG 850
838 ACTAGGCTCTTCACTGCGCTGTGACC 868

Db 851 ACCTA-GGGCTTTTCATCTGCCGCTGTGACC 880

RESULT 7
LOCUS B1768516
DEFINITION 603057118F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5206588 5', mRNA sequence.
ACCESSION B1768516
VERSION B1768516.1 GI:15760094
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 912)
NIH-MGC <http://mgc.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Plate: LLM1518 row: c column: 05
High quality sequence stop: 810.
Location/Qualifiers
1. 912
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5206588"
/clone_11b="NIH_MGC_122"
/lab_host="DH10B"
/note="Organ: pooled lung and spleen; Vector: PCMV-SPORT6; Site:1: NotI; Site:2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH-MGC Library."

BASE COUNT 230 a 262 c 225 g 195 t

Query Match 34.1% Score 793; DB 13; Length 912;
Best Local Similarity 96.1%; Pred. No. 5.3e-211;
Matches 856; Conservative 0; Mismatches 30; Indels 5; Gaps 4;

947 ATGAGTGTGTAACACGCGCGGACATCTTCTGCTCCGCTCAGGCTACATCTGC 1006
1 ATGAGTGTGTAACACGCGCGGACATCTTCTGCTCCGCTCAGGCTACATCTGC 60
1007 TGTGTGACCAACCGAAGCTGCCAAGACATCAAGATGTGAGACAGAACCACTGCA 1066
61 TGTGTGACCAACCGAAGCTGCCAAGACATCAAGATGTGAGACAGAACCACTGCA 120
1067 ACCTGACAGACAGCTGTACATTTTACAAAGGGGCTTCAATGATGACAGCCATCGCT 1126
121 ACCTGACAGACAGCTGTACATTTTACAAAGGGGCTTCAATGATGACAGCCATCGCT 180
1127 GTGAGGAGCTTATCTGAGGATCAAGTATTAACCGCTGTATGTCTGCTGAGAACCTTG 1186
181 GTGAGGAGCTTATCTGAGGATCAAGTATTAACCGCTGTATGTCTGCTGAGAACCTTG 240
1187 GCTGCAGAGACCAACCCCTTACATCTGTACCCGGGACATGAGACGCTGTGAGACGCT 1246
241 GCTGCAGAGACCAACCCCTTACATCTGTACCCGGGACATGAGACGCTGTGAGACGCT 300

REFERENCE
TITLE
AUTHORS
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 965)
NIH-MGC <http://mgc.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
<http://image.llnl.gov>
Plate: LLM13583 row: f column: 05
High quality sequence stop: 634.

FEATURES

source

1. 996
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5188164"
/clone_1lb="Lupski_sympathetic_trunk"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCAGCGCTCG-3' and
5'-GACTACTGATCGAGCGAGCGCCGCT(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."

BASE COUNT 245 a 284 c 242 g 224 t 1 others

ORIGIN

Query Match 35.5%; Score 825.4; DB 14; Length 996;

Best Local Similarity 97.9%; Pred. No. 4,4e-220;

Matches 889; Conservative 0; Mismatches 12; Indels 7; Gaps 5;

219 TCCAGCCCTGGGAATGACAGACAGTGCAGATGGCTTGACCTGATGCGCAGTC 278
1 TTCCAGCCCTGGGAATGACAGACAGTGCAGATGGCTTGACCTGATGCGCAGTC 60
279 AGGACAGTGTATGATATGATGATGCGGACACATCCCGAGGCTGCCGAGAGACAT 338
61 AGGACAGTGTATGATATGATGATGCGGACACATCCCGAGGCTGCCGAGAGACAT 120
339 GATGTGTGTTAACCAGAAATGGCGGTATTTATGATTCCTCCCGACAAACCTGTGTATCG 398
121 GATGTGTGTTAACCAGAAATGGCGGTATTTATGATTCCTCCCGACAAACCTGTGTATCG 180
339 AGGCGCTTACTGGAACCCCTACTGACACCCCTACTGATGATCCGATACAGAGTGCAGCC 458
181 AGGCGCTTACTGGAACCCCTACTGACACCCCTACTGATGATCCGATACAGAGTGCAGCC 240
459 ACCACTCTCAGCTCCAACTATCCACGATCTCCAGGCTCTTATATGCGCGCTTGTGATA 518
241 ACCACTCTCAGCTCCAACTATCCACGATCTCCAGGCTCTTATATGCGCGCTTGTGATA 300
519 CCAGTGTGATGAAGAACCAATGTGTGATGTGAGAGATGTGACACAGATTTCCACCA 578
301 CCAGTGTGATGAAGAACCAATGTGTGATGTGAGAGATGTGACACAGATTTCCACCA 360
579 GTGCAACCCACCCAGATCTGATCAATACTGAAAGGGGGGTACACGCTGCTCCAGACCA 638
361 GTGCAACCCACCCAGATCTGATCAATACTGAAAGGGGGGTACACGCTGCTCCAGACCA 420
639 CGGATATTTGGCTTCTGGAAGGCCAGTGTGATGACATTTGATGATGATGCTATGCTACTG 698

|||||
Db 421 CGGATATTTGGCTTCTGGAAGGCCAGTGTGATGACATTTGATGATGCTATGTTACTG 480
Oy 699 CCAGCAGCTGTGTGCAATGTTCTTCTGATCCTTATTTGTATGACAAACCTGTGTTAC 758
Db 481 CCGAGAGCTGTGTGCAATGTTCTTCTGATCCTTATTTGTATGACAAACCTGTGTTAC 540
Oy 759 CCGATATGAGATGGAAGGCTTGTGCAAGATGTGAGAGTGTGACAGACCCCG 818
Db 541 CCGATATGAGATGGAAGGCTTGTGCAAGATGTGAGAGTGTGACAGACCCCG 600
Oy 819 CCGTGAACCTGTGCTCAACACCTTACGCTTTCATGCTGCGCTGTGACCCAGATATGA 878
Db 601 CCGTGAACCTGTGCTCAACACCTTACGCTTTCATGCTGCGCTGTGACCCAGATATGA 660
Oy 879 ACTTGAGGAAGATGCGCTTATTCATGATGATGAGAGTGCAGCTTCTGTGACTTCT 938
Db 661 ACTTGAGGAAGATGCGCTTATTCATGATGATGAGAGTGCAGCTTCTGTGACTTCT 720
Oy 939 CCGCAACATGAGTGTGAAACCGCCGACATCTTGTCTCTGCGCC-TCGAGCT 997
Db 721 CCGCCACATGAGTGTGAAACCGCCGACATCTTGTCTCTGCGCC-TCGAGCT 780
Oy 998 ACATCTGCTGTGATGACACCGAAGCTGCGCAAGATCAAGATGTGAGACAGG-AAAC 1056
Db 781 ACATCTGCTGTGATGACACCGAAGCTGCGCAAGATCAAGATGTGAGACAGG-AAAC 840
Oy 1057 CACACGT-GCAACCTGACAGAG-ACGTCTCAATTT--ACAGGGGCTTCAATGCA 1111
Db 841 CACACGTGCGACCTGCGAGAGACGTCCTCAATTTACCAAGGGGCTTCAATGCA 900
Oy 1112 TCGACCC 1119
Db 901 TCGACCC 908

RESULT 6

LOCUS

DEFINITION

B1825121

603038662F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5179550 5',

ACCESSION

B1825121

VERSION

B1825121.1 GI:15936671

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

NIH-MGC <http://mgc.ncl.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

<http://image.llnl.gov>

Plate: LLM1447 row: 1 column: 15

High quality sequence start: 14

High quality sequence stop: 844.

FEATURES

source

1. 882
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5179550"
/clone_1lb="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
PCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA

Db 488 GACATGAGCTGTGTGTCAGACGCTCCGTCCTCCGTCACATCTTCCAAATGCAAGCCACG 547
 QY 1282 ACCGCTACCTGGGGGCTTATTTACATTTTCCAGATCAATCTGGGAATGAGGACAGAA 1341
 Db 548 ACCGCTACCTGGGGGCTTATTTACATTTTCCAGATCAATCTGGGAATGAGGACAGAA 607
 QY 1342 TTTTACATGCGGCACAAAGGGGCCCCATGACAGTCCCTGGTATGACAGGCCCATCAA 1401
 Db 608 TTTTACATGCGGCACAAAGGGGCCCCATGACAGTCCCTGGTATGACAGGCCCATCAA 667
 QY 1402 GGGGCCCCGGAAATCCAGCTGAGATTTGGAATGATCATCTGTAAACACTGTCACTCA 1461
 Db 668 GGGGCCCCGGAAATCCAGCTGAGATTTGGAATGATCATCTGTAAACACTGTCACTCA 727
 QY 1462 AGAGGACGCTCCGATCCGATCCGATCCGATATATGTGTGTCAGATACCATTTGAGGCTCG 1521
 Db 728 AGAGGACGCTCCGATCCGATCCGATCCGATATATGTGTGTCAGATACCATTTGAGGCTCG 787
 QY 1522 GCGTGGAGCTCCGAGCTGCTCTCTCATTTGGCACCAGGAGACAGAGAAGAAATA 1581
 Db 788 GCGTGGAGCTCCGAGCTGCTCTCTCATTTGGCACCAGGAGACAGAGAAGAAATA 847
 QY 1582 ACAGAGGATGAGAGGAGACACAGCTTAGCATTTCTGCTGAACCTTCCCGAATA 1641
 Db 848 ACCGAGGATGAGAGGAGACACAGCTTAGCATTTCTGCTGAACCTTCCCGAATA 907
 QY 1642 GTCAGCCCCGACCTTC-TGACTCTACCTGTACTATTGACAGACCTGT 1687
 Db 908 GTCAGCCCCGACCTTCCTTGACTCTCAGCTGTACTATTGACAGACCTGT 954

RESULT 4
 AL572040/c 876 bp mRNA linear EST 16-FEB-2001
 LOCUS AL572040 LTL.NFL006.PL2 Homo sapiens cDNA clone CS0D1026Y121 3
 DEFINITION prime, mRNA sequence.
 ACCESSION AL572040
 VERSION AL572040.1 GI:12929925
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 876)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 JOURNAL
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 FEATURES
 source
 1..876
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0D1026Y121"
 /clone_1id="LTL.NFL006.PL2"
 /tissue_type="placenta"
 /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifestech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 239 a 184 c 208 g 220 t 25 others
 ORIGIN

Query Match 36.0%; Score 837.4; DB 9; Length 876;
 Best Local Similarity 96.8%; Pred. No. 1.8e-223;

Matches 849; Conservative 1; Mismatches 26; Indels 1; Gaps 1;
 QY 1416 CCAGCTGACCTGGGAATGATCAGTCAACAGCTGTCAATCACTTACAGAGGACGCTCCGT 1475
 Db 876 CCAGCTGACCTGGGAATGATCAGTCAACAGCTGTCAATCACTTACAGAGGACGCTCCGT 817
 QY 1476 GATCCGACTGGGATATATGTGTGTCAGATACCATTTGAGGCTCGGGCTGAGGCTCCG 1535
 Db 816 GATCCGAGCNGGGANANANNGTNGCAGATACCCANNCTAGCGCTC-GGCTGGAGCNCGG 758
 QY 1536 ACCTGCTCTCATTTGGGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1595
 Db 757 ACCTGCTCTCATTTGGGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 698
 QY 1596 AGCAGACAGAGGTTAGGACATTTCTGCTGAACGTTTCCCGAAGAGTAGAGCCCGACTT 1655
 Db 697 AGCAGACAGAGGTTAGGACATTTCTGCTGAACGTTTCCCGAAGAGTAGAGCCCGACTT 638
 QY 1656 CCTGACTCTCAGCTGTACTATTGACAGACCTGTACCCCTGACAGACTTGCACCCGAGTT 1715
 Db 637 CCTGACTCTCAGCTGTACTATTGACAGACCTGTACCCCTGACAGACTTGCACCCGAGTT 578
 QY 1716 CCTATGATACAGTATATCAAAAAGTATATATCTGCTCCCTGATACAGATTTGTGTGA 1775
 Db 577 CCTATGACACAGTATATCAAAAAGTATATATCTGCTCCCTGATACAGATTTGTGTGA 518
 QY 1776 ATTTTCAAGGCTTCAGTTTATTTTCCACTATTTTCAAGAAATAGATTAGTTTGGGG 1835
 Db 517 ATTTTCAAGGCTTCAGTTTATTTTCCACTATTTTCAAGAAATAGATTAGTTTGGGG 458
 QY 1836 GGTCTGAGCTATGTTTCAAGAGCTGTGAACAGCTTGTCTCACTTCTTCACCTTTCAC 1895
 Db 457 GGTCTGAGCTATGTTTCAAGAGCTGTGAACAGCTTGTCTCACTTCTTCACCTTTCAC 398
 QY 1896 TCCCTTCTCAGCTGTGTACTGCTTGGCAAAACCCGGAGCTGGGGGGGAGACCTTGGGA 1955
 Db 397 TCCCTTCTCAGCTGTGTACTGCTTGGCAAAACCCGGAGCTGGGGGGGAGACCTTGGGA 338
 QY 1956 GTACCTAGTTTGTCTTGGCGTACAGAGAGGCTATGTAAACAACACACAGAGATC 2015
 Db 337 GTACCTAGTTTGTCTTGGCGTACAGAGAGGCTATGTAAACAACACACAGAGATC 278
 QY 2016 GAAGGGTTTTAGAGATGTGTTTCAAAACATGCTGTATTTTCAACCATTAAGAAG 2075
 Db 277 GAAGGGTTTTAGAGATGTGTTTCAAAACATGCTGTATTTTCAACCATTAAGAAG 218
 QY 2076 TTTAGTTTGTCTTAAATTTGATTAACGGTTTAAATTTCTGTTTCAATTTAGATTT 2135
 Db 217 TTTAGTTTGTCTTAAATTTGATTAACGGTTTAAATTTCTGTTTCAATTTAGATTT 158
 QY 2136 TTAATAAATATGTGTAGATTTCTTGAAGGCTTTCAGACATGCTATGTCTGTCT 2195
 Db 157 TTAATAAATATGTGTAGATTTCTTGAAGGCTTTCAGACATGCTATGTCTGTCTCT 98
 QY 2196 TCCCAAAACCAAGTCTCTCCATTTTAAAGCCAGTGTTCCTTGAAGACCCCTATCT 2255
 Db 97 TCCCAAAACCAAGTCTCTCCATTTTAAAGCCAGTGTTCCTTGAAGACCCCTATCT 38
 QY 2256 TGCCTTCTTGAATTTTACCAATTTGATTTGAAT 2292
 Db 37 TGCCTTCTTGAATTTTACCAATTTGATTTGAAT 1

RESULT 5
 LOCUS B0720516 996 bp mRNA linear EST 16-JUL-2002
 DEFINITION AGENCOURT_8219503 lupski_sympathetic_trunk Homo sapiens cDNA clone
 IMAGE:6188164 5', mRNA sequence.
 ACCESSION B0720516
 VERSION B0720516.1 GI:21859413
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

```

OY 210 TCTCTCTCTTCAAGCCCTGGGAATGCACAGGACAGTGCAGATGGCTTGGACCTGGA 269
    |||||
DB 129 TCTCTCTCTTCAAGCCCTGGGAATGCACAGGACAGTGCAGATGGCTTGGACCTGGA 188
OY 270 TCGCCAGTCAAGGACAGTGTATTGATATGATGAAATGCCAAACCATCCCGAGGCTGCCG 329
    |||||
DB 189 TCGCCAGTCAAGGACAGTGTATTGATATGATGAAATGCCAAACCATCCCGAGGCTGCCG 248
OY 330 AGGAGACATGATGTGTATTAACCAAAATGGCGGGGATTTATGATTCCTCCCGGCAAAACC 389
    |||||
DB 249 AGGAGACATGATGTGTATTAACCAAAATGGCGGGGATTTATGATTCCTCCCGGCAAAACC 308
OY 390 TGTGTATGAGAGGCGCTTACTGGAACCCCTACTGCAACCCCTACTGAGTTCCTGACCAAC 449
    |||||
DB 309 TGTGTATGAGAGGCGCTTACTGGAACCCCTACTGCAACCCCTACTGAGTTCCTGACCAAC 368
OY 450 AGCTGCCCCCACTCTCAGCTCCAAACTATCCACGATTCACGAGCTTTATATGCCG 509
    |||||
DB 369 AGCTGCCCCCACTCTCAGCTCCAAACTATCCACGATTCACGAGCTTTATATGCCG 428
OY 510 CTTTGATACCAATGGATGAAGAACCAACATGTGTGATGTGACGAGTGTGCAACAGA 569
    |||||
DB 429 CTTTGATACCAATGGATGAAGAACCAACATGTGTGATGTGACGAGTGTGCAACAGA 488
OY 570 TTCCCAACAGTGCACACCCCAACCCAGATCTGCATCAATCTGGAAGCGGGTACACCTGCTC 629
    |||||
DB 489 TTCCCAACAGTGCACACCCCAACCCAGATCTGCATCAATCTGGAAGCGGGTACACCTGCTC 548
OY 630 CTGACACGAGGATATTTGCTCTGGAAGCGGCACTGCTTGAACATTTGATGATGTCCTA 689
    |||||
DB 549 CTGACACGAGGATATTTGCTCTGGAAGCGGCACTGCTTGAACATTTGATGATGTCCTA 608
OY 690 TGGTTACTGACACAGCTGTGTGGAATGTTCCGATCCGATCTTCTTACATGGAACCC 749
    |||||
DB 609 TGGTTACTGACACAGCTGTGTGGAATGTTCCGATCCGATCTTCTTACATGGAACCC 668
OY 750 TGGTTTACCCTCAATGAGATGGAAGGTCTTGGCAAGATGGAACGAGTGGCCACGA 809
    |||||
DB 669 TGGTTTACCCTCAATGAGATGGAAGGTCTTGGCAAGATGGAACGAGTGGCCACGA 728
OY 810 GAAACCCCTGCGTCAAACTGCGTCAACACCTACGAGCTCTTTCATCTGCCGCTGACCC 869
    |||||
DB 729 GAAACCCCTGCGTCAAACTGCGTCAACACCTACGAGCTCTTTCATCTGCCGCTGACCC 788
OY 870 AGGATATGAATCTGAGAAATGGCTTTCATCTGAGTATGAGACGAGTGCAGCTTCTC 929
    |||||
DB 789 AGGATATGAATCTGAGAAATGGCTTTCATCTGAGTATGAGACGAGTGCAGCTTCTC 848
OY 930 TAGATTCCTCTGCAACATGATGTGTGAACCAACCCGACATCTCTGCTCTGCC 989
    |||||
DB 849 TAGATTCCTCTGCAACATGATGTGTGAACCAACCCGACATCTCTGCTCTGCC 908
OY 990 TCCAGGCTACATCTGCTGATGACAAACGAAAGTCCGCAAGACATCAAGATGTGAGA 1049
    |||||
DB 909 T-CAGGCTACATCTGCTGATGACAAACGAAAGTCCGCAAGATCAAGATGTGAGA 966
OY 1050 CAGGAAACCACTGCAACCTGCAAGAGAGTGTCTACATTTACAGGGGG 1100
    |||||
DB 967 CAGGAAACCACTGCAACCTGCAAGAGAGTGTCTACATTTACAGGGGG 1014

```

```

REFERENCE
AUTHORS 1 (Bases 1 to 954)
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LEM12727 row: c column: 05
High quality sequence stop: 676.

FEATURES
Source
Location/Qualifiers
1..954
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5729596"
/cisue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site:1; EcoRV
(destroyed); Site:2: Not; RNA source male hippocampus,
age 27. Library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.4 kb, insert size range 0.9-4 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 012."

BASE COUNT 232 a 278 c 238 g 206 t
ORIGIN
Query Match 39.8%; Score 925.4; DB 13; Length 954;
Best Local Similarity 99.3%; Pred. No. 3.8e-248;
Matches 940; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

OY 742 TGCACCCCTGTTTACCCTCAATGAGATGGAAGTCTTGGCAAGTGTGACGAGTGT 801
    |||||
DB 8 TGCACCCCTGTTTACCCTCAATGAGATGGAAGTCTTGGCAAGTGTGACGAGTGT 67
OY 802 GCCACGAGAACCCCTGGTGCACAACTGCGGCAACACCTACGGCTCTTCATCTGCCG 861
    |||||
DB 68 GCCACGAGAACCCCTGGTGCACAACTGCGGCAACACCTACGGCTCTTCATCTGCCG 127
OY 862 TGTGACCCAGATATGAACTTGAAGATGAGCGCTTCAATTGACGATATGACGAGTGC 921
    |||||
DB 128 TGTGACCCAGATATGAACTTGAAGATGAGCGCTTCAATTGACGATATGACGAGTGC 187
OY 922 AGCTTCTGTAGTCTCTGCAACATGATGTGTGAACCAACCCGACATCTTCTGCG 981
    |||||
DB 188 AGCTTCTGTAGTCTCTGCAACATGATGTGTGAACCAACCCGACATCTTCTGCG 247
OY 982 TCCGCTCCAGGCTACATCTGCTGATGACAAACGAAAGTGGCAAGACATCAACGAA 1041
    |||||
DB 248 TCCGCTCCAGGCTACATCTGCTGATGACAAACGAAAGTGGCAAGACATCAACGAA 307
OY 1042 TGTGAGCAGAGAACCAACAGCTGCAACCTGCAAGAGAGTGTCTACATTTACAGGGGG 1101
    |||||
DB 308 TGTGAGCAGAGAACCAACAGCTGCAACCTGCAAGAGAGTGTCTACATTTACAGGGGG 367
OY 1102 TTCAAAATGATGCAACCCCATCCGCTGTGAGAGACCTTATCTGAGATCATGATTAACCCG 1161
    |||||
DB 368 TTCAAAATGATGCAACCCCATCCGCTGTGAGAGACCTTATCTGAGATCATGATTAACCCG 427
OY 1162 TGTATGTCTCTGCTGGAACCCGCTGCAAGAGAGACCAAGCCCTTACATCTTGACCGG 1221
    |||||
DB 428 TGTATGTCTCTGCTGGAACCCGCTGCAAGAGAGACCAAGCCCTTACATCTTGACCGG 487
OY 1222 GACATGACGTGTGCTGAGAGAGCTCCGCTGCAATCTTCAATGACAGCCAGG 1281
    |||||

```

RESULT 3
 LOCUS BM545444 954 bp mRNA linear EST 20-FEB-2002
 DEFINITION AGENCOURT_6500472 NIH_MGC_124 Homo sapiens CDNA clone IMAGE:5729596
 5', mRNA sequence.
 ACCESSION BM545444
 VERSION BM545444.1 GI:18777513
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 293 a 224 c 262 g 268 t 6 others

Query Match 44.8% Score 1042.2; DB 9; Length 1053;
Best Local Similarity 99.0%; Pred. No. 6.6e-281;
Matches 1042; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

```

1242 ACCGTCGTCCTCCGCTGACATCTCCAAATGCAAGCCGACCCCTACCTGGGCGCTA 1301
1053 ACCGTCGTCCTCCGCTGACATCTCCAAATGCAAGCCGACCCCTACCTGGGCGCTA 994
1302 TTACATTTTCCAGATCAAAATCTGGGAATGAGGCGAGAAATTTTACATCGGCAACGGG 1361
993 TTACATTTTCCAGATCAAAATCTGGGAATGAGGCGAGAAATTTTACATCGGCAACGGG 934
1362 CCCCATCAGTCCACCTCGTGATGACACGCCCATCAAAAGGCCCCGGGAATCCAGCT 1421
933 CCCCATCAGTCCACCTCGTGATGACACGCCCATCAAAAGGCCCCGGGAATCCAGCT 874
1422 GGACCTGGGAATGATCACTGCAACATGTCATCAATTCAGAGGACAGCTCGGTATCCG 1481
873 GGACCTGGGAATGATCACTGCAACATGTCATCAATTCAGAGGACAGCTCGGTATCCG 814
1482 ACTGCGGATATATGTGTGCGAGTACCCATTCGAGCTCGGGCTGAGCTCGAGCTG 1541
813 ACTGCGGATATATGTGTGCGAGTACCCATTCGAGCTCGGGCTGAGCTCGAGCTG 754
1542 CCTCTCTTTGGCACCAGGAGCAGAGAGAGAAATTAACAGAGAGATGAGAGCGAC 1601
753 CCTCTCTTTGGCACCAGGAGCAGAGAGAGAAATTAACAGAGAGATGAGAGCGAC 694
1602 ACAGAGCTTGAAGCTTTCCTGTAAGCTTCCCGGAAGAGTACAGCCCGGATTCCTGAC 1661
693 ACAGAGCTTGAAGCTTTCCTGTAAGCTTCCCGGAAGAGTACAGCCCGGATTCCTGAC 634
1662 TCTCAGCTGATATTCAGACCTGTCAACCTGACAGACTTCCACCCCGATTCCTATG 1721
633 TCTCAGCTGATATTCAGACCTGTCAACCTGACAGACTTCCACCCCGATTCCTATG 574
1722 ATACAGTATCAAAATATATTCATGCTCCCTCGATAGAGATTTGGTGAATTTTC 1781
573 ACACAGTATCAAAATATATTCATGTAACCCCGATAGAGATTTGGTGAATTTTC 514
1782 AAGGCTTCAGTTATTTCCACTATTTTCAAGAAATATAGATTAGTTGGGGGCTCG 1841
513 AAGGCTTCAGTTATTTCCACTATTTTCAAGAAATATAGATTAGTTGGGGGCTCG 454
1842 AGCTATATGTTCAAAAGTGTGAACAGCTTGTCTACCTTTCACCTTCACCTCTTC 1901
453 AGCTATATGTTCAAAAGTGTGAACAGCTTGTCTACCTTTCACCTTCACCTCTTC 394
1902 TCTCAGCTGATATTCAGACCTGTCAACCTGACAGACTTCCACCCCGATTCCTATG 1961
393 TCTCAGCTGATATTCAGACCTGTCAACCTGACAGACTTCCACCCCGATTCCTATG 334
1962 AGTTTCCTTTTGGTACAGAGAGAGCTATGTAAACAAACACAGCAGATCGAAGG 2021
333 AGTTTCCTTTTGGTACAGAGAGAGCTATGTAAACAAACACAGCAGATCGAAGG 274
2022 TTTTATGAGAAATGTTTCAAAACCATGCTGTATTTTCAACCATATAAAGAGTTTCAG 2081
273 TTTTATGAGAAATGTTTCAAAACCATGCTGTATTTTCAACCATATAAAGAGTTTCAG 214
2082 TTGTCTTAATTTGTTAAGGTTTAAATCTGTCTTCTTCAATTTTGAATTTTAA 2141

```

```

Db 213 TTGTCTTAATTTGTTAAGGTTTAAATCTGTCTTCTTCAATTTTGAATTTTAA 154
QY 2142 AATATGCTGTAATTCCTTGAAGGCTTACAGACATGCTATGCTGCTCCCA 2201
Db 153 AATATGCTGTAATTCCTTGAAGGCTTACAGACATGCTATGCTGCTCCCA 94
QY 2202 ACCAGAGTCTCTTCATTTTACCCAGTGTCTTTTGGAGACCCCTTAATCTGCTT 2261
Db 93 ACCAGAGTCTCTTCATTTTACCCAGTGTCTTTTGGAGACCCCTTAATCTGCTT 34
QY 2262 CTTTGAATTTTACCCAGTGTGTGATGATGATGATGATGATGATGATGATGATG 2294
Db 33 CTTTGAATTTTACCCAGTGTGTGATGATGATGATGATGATGATGATGATGATGATG 1

RESULT 2
BM921371 1014 bp mRNA linear EST 13-MAR-2002
LOCUS BM921371
DEFINITION AGENCOURT_6626255 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5752785
5', mRNA sequence.
ACCESSION BM921371
VERSION BM921371.1 GI:19371750
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 1014)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rstraus@nhi.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
plate: LLM1287 row: 1 column: 10
High quality sequence stop: 738.
Location/Qualifiers
1. 1014
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5752785"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site: 1: NotI; Site: 2: EcoRV (destroyed); RNA
source: anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research genetics tracking code
021. Note: this is a NIH_MGC Library."
BASE COUNT 235 a 300 c 246 g 233 t
ORIGIN
Query Match 40.5%; Score 943.4; DB 14; Length 1014;
Best Local Similarity 98.4%; Pred. No. 3.4e-253;
Matches 995; Conservative 0; Mismatches 11; Indels 5; Gaps 4;
QY 90 GAGCTGCCGAGCTCTCCGCGGCTCCGCGCGAGCTTCTCTGCGCTTGCATCTC 149
Db 9 GATGTGCCCGAGCTCTCCGCGGCTCCGCGCGAGCTTCTCTGCGCTTGCATCTC 68
QY 150 CTCTCGCGGCTCTGGACATGCCAGAAATAAAGATACACATGTTACATCTGAC 209
Db 69 CTCTCGCGGCTCTGGACATGCCAGAAATAAAGATACATGTTACATCTGAC 128

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 10:48:06 ; Search time 3433.81 Seconds
(Without alignments)
10979.953 Million cell updates/sec

Title: US-09-674-379A-12
Perfect score: 2328
Sequence: 1 gaccgcgcgcctcccccgtg.....tgatataatttgaagaga 2328

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARY

Result No.	Score	Query Match length	ID	Description
C 1	1042.2	44.8	1053 9	AL571218 AL571218
2	943.4	40.5	1014 14	BM921371 AGENCOURT
3	925.4	39.8	954 13	BM545444 AGENCOURT
4	837.4	36.0	876 9	AL572040 AL572040
5	825.4	35.5	996 14	BQ720516 AGENCOURT
6	807.6	34.7	882 13	BI825121 BI825121

7	793	34.1	912 13	BI768516
8	792.2	34.0	904 14	BQ716569
9	790.6	34.0	880 13	BI764020
10	785	33.7	835 14	BQ718885
C 11	723.8	31.1	727 9	AL570066
C 12	719.4	30.9	721 13	BI523975
C 13	717.4	30.8	735 14	BM974454
C 14	700.4	30.1	794 13	BI837271
15	700.4	30.1	1006 9	AL545549
C 16	682.4	29.7	709 14	BQ015079
C 17	683.4	29.4	699 14	BM970551
C 18	674.6	29.0	762 13	BI834712
19	666.8	28.6	891 12	BF339448
C 20	665	28.6	822 12	BI109391
C 21	650.4	27.9	667 14	BM970136
C 22	650.4	27.9	669 13	BM666348
C 23	646.4	27.8	889 14	BQ897144
24	639.2	27.5	796 13	BI453348
25	637.6	27.4	870 14	BQ930927
26	636	27.3	644 13	BG938621
C 27	630.4	27.1	672 14	BQ006021
C 28	626	26.9	803 12	BE786513
29	620.6	26.7	885 9	AL543659
C 30	615.2	26.4	1015 13	BM544363
C 31	608.4	26.1	627 14	BQ772930
C 32	606	26.0	629 13	BI489863
33	600.2	25.8	906 14	BQ723059
C 34	592.8	25.5	717 13	BI251219
35	584.6	25.1	886 12	BG485593
36	576.8	24.8	690 12	BF104873
37	576.4	24.8	933 14	BQ716088
38	575.8	24.7	770 12	BQ542457
C 39	573.6	24.6	627 12	BF090699
C 40	573	24.5	573 10	AM183044
41	571	24.5	587 14	BM690459
42	570.2	24.5	586 14	BM695168
C 43	568	24.4	586 13	BG938622
C 44	565.6	24.3	572 10	AM299380
45	564.8	24.3	815 12	BG174360

ALIGNMENTS

RESULT 1
LOCUS AL571218/c 1053 bp mRNA linear EST 16-FEB-2001
DEFINITION LTI_NFL006_PL2 Homo sapiens CDNA clone CS001015YE20 3
ACCESSION AL571218
VERSION AL571218.1 GI:12928294
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1053)
AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

1..1053
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS001015YE20"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand CDNA
was primed with a NotI-oligo(dT) primer. Five prime end

XX (ONCY) ONO PHARM CO LTD.
PA
XX

PI Honjo T, Tashiro K, Nakamura T;
XX

DR WPI: 2000-038646/03.
XX

DR N-PSDB; AA239385.
XX

PT Novel mouse polypeptides for treatment, e.g. arteriosclerosis and myoma
XX

PS Claim 1; Page 51-53; 70pp; Japanese.
XX

CC The invention provides mouse polypeptides for treatment of diseases due
CC to abnormal proliferation of smooth muscle. The polypeptides can be
CC produced by standard recombinant methodology. The polypeptides can be
CC used according to their inhibition of the proliferation of vascular
CC smooth muscle cells, particularly in treating arteriosclerosis or re-
CC narrowing by vascular endothelial thickening after percutaneous
CC transluminal coronary angioplasty (PTCA), or myoma, hematopoietic cell-
CC regulatory activity, cytokine activity, tissue generation/repatriation
CC activity, actin/inhibitor activity, taxis and chemotaxis activity, blood
CC coagulation/thrombotic activity, receptor/ligand activity, cadherin/
CC tumour metastasis inhibiting activity; tumor inhibition, and as nutrient.
CC The present sequence represents the mature protein of the invention which
CC can be used for modulating smooth muscle cell proliferation.
XX

SQ Sequence 423 AA:

Query Match 90.9%; Score 2302; DB 21; Length 423;
Best Local Similarity 94.8%; Pred. No. 1.4e-148;

Matches 401; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

QY 26 OCTNGFDLDROSGGQCLDIDECRTIPACRGDMCVNONGYLCTPRTNPTVYRGPSNPS 85
DB 1 OCTNGFDLDROSGGQCLDIDECRTIPACRGDMCVNONGYLCTPRTNPTVYRGPSNPS 60
QY 86 TPYSGPYPAAPPLASAPNTPTISPLICRFGYQMDSESNQCVDECATDSHCNPTQICI 145
DB 61 TSYSGPYPAAPPLASAPNTPTISPLICRFGYQMDSESNQCVDECATDSHCNPTQICI 120
QY 146 NTEGGYTCSTDGWMLLEGGQCLDIDECRYGYCOQLCANVPQSYSCGCPGFTLNEDGRSC 205
DB 121 NTEGGYTCSTDGWMLLEGGQCLDIDECRYGYCOQLCANVPQSYSCGCPGFTLNEDGRSC 180
QY 206 ODVNECATENPCVQTCVNTYSGFICRCDPGYLEEDGVHCSMDCESEFLCOHECVNQ 265
DB 181 ODVNECATENPCVQTCVNTYSGFICRCDPGYLEEDGVHCSMDCESEFLCOHECVNQ 240
QY 266 PGTYFCSCPPIGLDLDNRSCQDINECEHNHNTCNLOQTCYNLOGGFRCIDPICEEPTL 325
DB 241 PGTYFCSCPPIGLDLDNRSCQDINECEHNHNTCNLOQTCYNLOGGFRCIDPICEEPTL 300
QY 326 RISDNRCMCPAENPGCRDQPTLLYRDMVDVSGRSVPADIFQMATRYPGAYYIFQIKS 385
DB 301 LISDNRCMCPAENPGCRDQPTLLYRDMVDVSGRSVPADIFQMATRYPGAYYIFQIKS 360
QY 386 GNEGRFVYRQGTISATLMTPIKGRRIOLDLEMITVNTVINFGRSSVIRLRIYYSQ 445
DB 361 GNEGRFVYRQGTISATLMTPIKGRRIOLDLEMITVNTVINFGRSSVIRLRIYYSQ 420
QY 446 YPF 448
DB 421 YPF 423

Search completed: July 3, 2003, 17:54:48
Job time : 34.4041 secs


```

QY 65 GYLCTPRTNPVYRGPSNPSTYSGPYPAAPPLSAPNYPRTSRPLICRFGYOMDSNQ 124
DB 78 GYLCTPRTNPVYRGPSNPSTYSGPYPAAPPLSAPNYPRTSRPLICRFGYOMDSNQ 137
QY 125 CVDVDECATDSSHQCNPQTQICINTEGGYTCSTDGWYLLLEGQCLDIDECRYGYCOQLCANV 184
DB 138 CVDVDECATDSSHQCNPQTQICINTEGGYTCSTDGWYLLLEGQCLDIDECRYGYCOQLCANV 197
QY 185 PGSYCTCNPGFTLNDGRCODVNECATENPCVQVCVNTYGSFTICRCDGYLEEGVH 244
DB 198 PGSYCTCNPGFTLNDGRCODVNECATENPCVQVCVNTYGSFTICRCDGYLEEGVH 257
QY 245 CSDMDECSFSEFLCOHECVNPGTYFCSCPGYIILLDNRSODINECHRNHTCNLQOT 304
DB 258 CSDMDECSFSEFLCOHECVNPGTYFCSCPGYIILLDNRSODINECHRNHTCNLQOT 317
QY 305 CYNLOGGFKCIDPIRCEEPYLRISDNRCMCPAENPCRDQPFITLYRDMDVSGRSVPAD 364
DB 318 CYNLOGGFKCIDPIRCEEPYLRISDNRCMCPAENPCRDQPFITLYRDMDVSGRSVPAD 377
QY 365 IFOMATRRYPGAYYIFQIKSGNEGREFYMRQTPISATLYMTRPIKGPREDIDLEMIT 424
DB 378 IFOMATRRYPGAYYIFQIKSGNEGREFYMRQTPISATLYMTRPIKGPREDIDLEMIT 437
QY 425 VNTVINFGSSSVIRLRITYSOYPF 448
DB 438 VNTVINFGSSSVIRLRITYSOYPF 461

RESULT 14
AAV54991
ID AAV54991 standard; Protein; 461 AA.
XX
AC AAV54991;
DT 15-FEB-2000 (first entry)
DE Full length mouse A55 protein sequence.
XX
KW A55 protein; mouse; smooth muscle proliferation; tissue generation;
KW vascular smooth muscle cell; arteriosclerosis; tissue reparation; myoma;
KW percutaneous transluminal coronary angioplasty; blood coagulation; PRCA;
KW actin; inhibin; chemotaxis; thrombosis; cadherin; therapy;
KW tumour metastasis inhibitor; ss.
XX
XX Mus musculus.
XX
XX WO9955864-A1.
XX
XX 04-NOV-1999.
XX
XX 28-APR-1999; 99WO-JP02284.
XX
XX 28-APR-1998; 98JP-0119731.
XX
XX (ONOV ) ONO PHARM CO LTD.
XX
XX Honjo T, Tashiro K, Nakamura T;
XX
XX WPI: 2000-038647/03.
XX
XX N-PSDB; AAZ40031.
XX
XX Novel human polypeptides for treatment of, e.g. arteriosclerosis and
XX myoma -
XX
XX Example 5; Page 64-68; 87pp; Japanese.
XX
XX This sequence is the mouse A55 protein. The invention relates to the
XX human A55 protein. The protein can be used for the treatment of diseases
XX due to abnormal proliferation of smooth muscle. The polypeptides can be
XX used according their inhibition of the proliferation of vascular smooth

```

```

CC muscle cells, particularly in treating arteriosclerosis or re-narrowing
CC by vascular endothelial thickening after percutaneous transluminal
CC coronary angioplasty (PRCA), or myoma, hematopoietic cell-regulatory
CC activity, cytokine activity, tissue generation/reparation activity,
CC actin/inhibin activity, taxis and chemotaxis activity, blood
CC coagulation/thrombotic activity, receptor/ligand activity,
CC cadherin/tumour metastasis inhibiting activity; tumour inhibition, and as
CC a nutrient.
XX
XX Sequence 461 AA:
XX
XX Query Match 93.8%; Score 2376; DB 21; Length 461;
XX Best Local Similarity 93.9%; Pred. No. 1.5e-153;
XX Matches 417; Conservative 10; Mismatches 17; Indels 0; Gaps 0;
XX
QY 5 KRILVTITALLCLPBGNAOACTNFGDIDRSGGCLDIDECRTIPACRGDMVCVQNG 64
DB 18 KRILVTITALLCLPBGNAOACTNFGDIDRSGGCLDIDECRTIPACRGDMVCVQNG 77
QY 65 GYLCTPRTNPVYRGPSNPSTYSGPYPAAPPLSAPNYPRTSRPLICRFGYOMDSNQ 124
DB 78 GYLCTPRTNPVYRGPSNPSTYSGPYPAAPPLSAPNYPRTSRPLICRFGYOMDSNQ 137
QY 125 CVDVDECATDSSHQCNPQTQICINTEGGYTCSTDGWYLLLEGQCLDIDECRYGYCOQLCANV 184
DB 138 CVDVDECATDSSHQCNPQTQICINTEGGYTCSTDGWYLLLEGQCLDIDECRYGYCOQLCANV 197
QY 185 PGSYCTCNPGFTLNDGRCODVNECATENPCVQVCVNTYGSFTICRCDGYLEEGVH 244
DB 198 PGSYCTCNPGFTLNDGRCODVNECATENPCVQVCVNTYGSFTICRCDGYLEEGVH 257
QY 245 CSDMDECSFSEFLCOHECVNPGTYFCSCPGYIILLDNRSODINECHRNHTCNLQOT 304
DB 258 CSDMDECSFSEFLCOHECVNPGTYFCSCPGYIILLDNRSODINECHRNHTCNLQOT 317
QY 305 CYNLOGGFKCIDPIRCEEPYLRISDNRCMCPAENPCRDQPFITLYRDMDVSGRSVPAD 364
DB 318 CYNLOGGFKCIDPIRCEEPYLRISDNRCMCPAENPCRDQPFITLYRDMDVSGRSVPAD 377
QY 365 IFOMATRRYPGAYYIFQIKSGNEGREFYMRQTPISATLYMTRPIKGPREDIDLEMIT 424
DB 378 IFOMATRRYPGAYYIFQIKSGNEGREFYMRQTPISATLYMTRPIKGPREDIDLEMIT 437
QY 425 VNTVINFGSSSVIRLRITYSOYPF 448
DB 438 VNTVINFGSSSVIRLRITYSOYPF 461

RESULT 15
AAV56751
ID AAV56751 standard; Protein; 423 AA.
XX
XX AAV56751;
XX
XX 22-FEB-2000 (first entry)
XX
XX Smooth muscle proliferation modulating protein mature sequence.
XX
XX Cell proliferation; vascular; smooth muscle cell; arteriosclerosis; PRCA;
XX endothelial thickening; percutaneous transluminal coronary angioplasty;
XX myoma; hematopoietic cell-regulation; cytokine; tissue generation; taxis;
XX actin activity; chemotaxis; blood coagulation; thrombotic; tumor; mouse;
XX metastasis; nutrient.
XX
XX Mus musculus.
XX
XX WO9955863-A1.
XX
XX 04-NOV-1999.
XX
XX 28-APR-1999; 99WO-JP02283.
XX
XX 28-APR-1998; 98JP-0119731.

```


CC can be modulated with agents that bind to these polypeptides, resulting
 CC in the death of the cells. The polynucleotides encoding these
 CC polypeptides are useful in the recombinant production of the
 CC polypeptides, as a hybridisation probe to screen libraries to isolate
 CC homologous sequences, or to map the gene. They may also be used for
 CC analysing genetic disorders, and to produce transgenic animals which are
 CC useful for the development and screening of therapeutically useful
 CC reagents. The polynucleotides can also be used in gene therapy e.g. to
 CC replace a defective gene.

SO Sequence 448 AA:

Query Match 99.8%; Score 2527; DB 22; Length 448;
 Best Local Similarity 99.8%; Pred. No. 8.1e-164;
 Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPGKRLITVITLALCLPSGNAOCTNGFPLDRSGOCLDIDECRTPEACRGDMCV 60
 Db 1 MPGKRLITVITLALCLPSGNAOCTNGFPLDRSGOCLDIDECRTPEACRGDMCV 60
 QY 61 NONGYLCIPRTNPNVYRGPSNPNSTPYSGPYPAAPPLSAPNPTISRLICRFGYQMD 120
 Db 61 NONGYLCIPRTNPNVYRGPSNPNSTPYSGPYPAAPPLSAPNPTISRLICRFGYQMD 120
 QY 121 ESNOCVVDDECATSHOCCNPQICINTEGGYCTSDGYWLLGQCLDIDECRYGCOOL 180
 Db 121 ESNOCVVDDECATSHOCCNPQICINTEGGYCTSDGYWLLGQCLDIDECRYGCOOL 180
 QY 181 CANPGSYCTCNPGFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGYELE 240
 Db 181 CANPGSYCTCNPGFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGYELE 240
 QY 241 DGVHCSMDDECSFSEFLCOHECVNPGTYFCSCPPGYILLDNRSCODINECHRNHTCN 300
 Db 241 DGVHCSMDDECSFSEFLCOHECVNPGTYFCSCPPGYILLDNRSCODINECHRNHTCN 300
 QY 301 LQOTCYNLQGGFKCIDPICEPEYLRISDNRCMPAENPCGRDQPTIILYRDMVYSGRS 360
 Db 301 LQOTCYNLQGGFKCIDPICEPEYLRISDNRCMPAENPCGRDQPTIILYRDMVYSGRS 360
 QY 361 VPADIFOMQATTRYPGAYYIFQIKSGNEGREFFYMRGTGPISATLVMTPIKGPRIQDL 420
 Db 361 VPADIFOMQATTRYPGAYYIFQIKSGNEGREFFYMRGTGPISATLVMTPIKGPRIQDL 420
 QY 421 EMITVNTVINFRGSSVIRLRIVYSQYPF 448
 Db 421 EMITVNTVINFRGSSVIRLRIVYSQYPF 448

RESULT 11

AA56750
 ID AA56750 standard; Protein: 448 AA.

AA56750:

22-FEB-2000 (first entry)

Smooth muscle proliferation modulating protein.

CC Cell proliferation; vascular; smooth muscle cell; arteriosclerosis; PTCA;
 CC endothelial thickening; percutaneous transluminal coronary angioplasty;
 CC myoma; hematopoietic cell-regulation; cytokine; tissue generation; taxis;
 CC actin activity; chemotaxis; blood coagulation; thrombotic; tumor; mouse;
 CC metastasis; nutrient.

XX Mus musculus.

XX W09955863-A1.

XX 04-NOV-1999.

XX 28-APR-1999; 99WO-JP02283.

XX

PR 28-APR-1998; 98JP-0119731.

XX (ONOX) ONO PHARM CO LTD.

XX Honjo T, Tashiro K, Nakamura T;

XX WPI: 2000-038646/03.

XX DR N-PSDB; AA239383, AA239384.

XX Novel mouse polypeptides for treatment, e.g. arteriosclerosis and myoma

XX Claim 1; Page 42-44; 70pp; Japanese.

CC The invention provides mouse polypeptides for treatment of diseases due
 CC to abnormal proliferation of smooth muscle. The polypeptides can be
 CC produced by standard recombinant methodology. The polypeptides can be
 CC used according to their inhibition of the proliferation of vascular
 CC smooth muscle cells, particularly in treating arteriosclerosis or re-
 CC narrowing by vascular endothelial thickening after percutaneous
 CC transluminal coronary angioplasty (PTCA), or myoma, hematopoietic cell-
 CC regulatory activity, cytokine activity, tissue generation/repairation
 CC activity, actin/inhibitor activity, taxis and chemotaxis activity, blood
 CC coagulation/chromobiotic activity, receptor/ligand activity, cadherin/
 CC tumour metastasis inhibiting activity; tumor inhibition, and as nutrient.
 CC The present sequence represents the protein of the invention which can be
 CC used for modulating smooth muscle cell proliferation.

SO Sequence 448 AA:

Query Match 94.9%; Score 2405; DB 21; Length 448;
 Best Local Similarity 94.2%; Pred. No. 1.6e-155;
 Matches 422; Conservative 10; Mismatches 16; Indels 0; Gaps 0;

QY 1 MPGKRLITVITLALCLPSGNAOCTNGFPLDRSGOCLDIDECRTPEACRGDMCV 60
 Db 1 MPGKRLITVITLALCLPSGNAOCTNGFPLDRSGOCLDIDECRTPEACRGDMCV 60
 QY 61 NONGYLCIPRTNPNVYRGPSNPNSTPYSGPYPAAPPLSAPNPTISRLICRFGYQMD 120
 Db 61 NONGYLCIPRTNPNVYRGPSNPNSTPYSGPYPAAPPLSAPNPTISRLICRFGYQMD 120
 QY 121 ESNOCVVDDECATSHOCCNPQICINTEGGYCTSDGYWLLGQCLDIDECRYGCOOL 180
 Db 121 ESNOCVVDDECATSHOCCNPQICINTEGGYCTSDGYWLLGQCLDIDECRYGCOOL 180
 QY 181 CANPGSYCTCNPGFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGYELE 240
 Db 181 CANPGSYCTCNPGFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGYELE 240
 QY 241 DGVHCSMDDECSFSEFLCOHECVNPGTYFCSCPPGYILLDNRSCODINECHRNHTCN 300
 Db 241 DGVHCSMDDECSFSEFLCOHECVNPGTYFCSCPPGYILLDNRSCODINECHRNHTCN 300
 QY 301 LQOTCYNLQGGFKCIDPICEPEYLRISDNRCMPAENPCGRDQPTIILYRDMVYSGRS 360
 Db 301 LQOTCYNLQGGFKCIDPICEPEYLRISDNRCMPAENPCGRDQPTIILYRDMVYSGRS 360
 QY 361 VPADIFOMQATTRYPGAYYIFQIKSGNEGREFFYMRGTGPISATLVMTPIKGPRIQDL 420
 Db 361 VPADIFOMQATTRYPGAYYIFQIKSGNEGREFFYMRGTGPISATLVMTPIKGPRIQDL 420
 QY 421 EMITVNTVINFRGSSVIRLRIVYSQYPF 448
 Db 421 EMITVNTVINFRGSSVIRLRIVYSQYPF 448

RESULT 12

AA54990

ID AA54990 standard; Protein: 448 AA.

XX AA54990;

XX

CC the proliferation or differentiation of chondrocyte cells. The PRO
 CC proteins can be used to determine the presence of tumours and also
 CC susceptibility to tumour development, particularly adrenal, lung, colon,
 CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
 CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
 CC can be used for genetic analysis of individuals with genetic disorders.

XX Sequence 448 AA;

Query Match 99.88; Score 2527; DB 22; Length 448;
 Best Local Similarity 99.88; Pred. No. 8,1e-164;
 Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGIKRLVITIALCLPSPGNAQAQCTNGFDLRQSGQCLDIDECRTIPEACRGDMCV 60
 DB 1 MGIKRLVITIALCLPSPGNAQAQCTNGFDLRQSGQCLDIDECRTIPEACRGDMCV 60
 QY 61 NONGGYLCIRTNVYRGPSNPTSPYSGPYPAAPPLSAPNPTISNPLICRGYQMD 120
 DB 61 NONGGYLCIRTNVYRGPSNPTSPYSGPYPAAPPLSAPNPTISNPLICRGYQMD 120
 QY 121 ESNQCVVDECATDSDHOCNPTQICINTEGCTCTDGYWLLGQCLDIDECRYGCOOL 180
 DB 121 ESNQCVVDECATDSDHOCNPTQICINTEGCTCTDGYWLLGQCLDIDECRYGCOOL 180
 QY 181 CANVPGSYCTCNGFTLNEDEGRSCQDVNECATENPCVOTCVNTYGSFICRDPGELEE 240
 DB 181 CANVPGSYCTCNGFTLNEDEGRSCQDVNECATENPCVOTCVNTYGSFICRDPGELEE 240
 QY 241 DGVHCSMDSCSFSEFLCQHCVCVNOPTGYFSCSPGGYILLDNRSCODINECHRNHTN 300
 DB 241 DGVHCSMDSCSFSEFLCQHCVCVNOPTGYFSCSPGGYILLDNRSCODINECHRNHTN 300
 QY 301 LQOTCYNIAGGFKCIDPRCEPFLRISDNRCMCAENPGCDOPFTIILYRMDVYSGRS 360
 DB 301 LQOTCYNIAGGFKCIDPRCEPFLRISDNRCMCAENPGCDOPFTIILYRMDVYSGRS 360
 QY 361 VPADIFOMQATTRYPGAYYIFQIKSGNEGRFYMROTGPISATLVWTRPIKGPRIQLDL 420
 DB 361 VPADIFOMQATTRYPGAYYIFQIKSGNEGRFYMROTGPISATLVWTRPIKGPRIQLDL 420
 QY 421 EMITVNTVINRGSSVRLRIYVSOYPP 448
 DB 421 EMITVNTVINRGSSVRLRIYVSOYPP 448

RESULT 10

ID AAB31183 standard; Protein: 448 AA.

XX AAB31183;

DT 20-APR-2001 (first entry)

XX Amino acid sequence of human polypeptide PRO210.

XX Human; secreted protein; transmembrane protein; PRO196; PRO444; PRO183;
 KW PRO185; PRO210; PRO215; PRO242; PRO288; PRO365; PRO361; PRO1308;
 KW PRO1183; PRO1272; PRO1419; PRO4999; PRO7170; PRO248; PRO353; PRO1318;
 KW PRO1600; PRO9940; PRO533; PRO301; PRO187; PRO337; PRO4111; PRO4356;
 KW PRO246; PRO265; PRO941; PRO10096; PRO6003; PRO6004; PRO2630;
 KW PRO6309; cell death; genetic disorder; transgenic animal; gene therapy.

XX Homo sapiens.

OS Key Location/Qualifiers

FT Peptide 1..25 /note- "signal peptide"

FT Modified-site 21..27 /note- "N-myristoylation site"

FT Binding-site 54..57 /note- "cell attachment site"

FT Modified-site 64..70

FT /note- "N-myristoylation site"
 FT 144..156
 FT /note- "aspartic acid and asparagine hydroxylation site"
 FT 149..155
 FT /note- "N-myristoylation site"
 FT 186..192
 FT /note- "N-myristoylation site"
 FT 226..232
 FT /note- "N-myristoylation site"
 FT 242..248
 FT /note- "N-myristoylation site"
 FT 267..273
 FT /note- "N-myristoylation site"
 FT 283..287
 FT /note- "N-glycosylation site"
 FT 296..300
 FT /note- "N-glycosylation site"
 FT 310..316
 FT /note- "N-myristoylation site"

WO20007037-A2.

PD 21-DEC-2000.

PE 22-MAY-2000; 2000WO-US14042.

XX 15-JUN-1999; 99US-0139695.
 PR 20-JUL-1999; 99US-0145070.
 PR 26-JUL-1999; 99US-0145698.
 PR 17-AUG-1999; 99US-0149396.
 PR 01-SEP-1999; 99WO-US20111.
 PR 08-SEP-1999; 99WO-US20594.
 PR 15-SEP-1999; 99WO-US21090.
 PR 30-NOV-1999; 99WO-US28313.
 PR 01-DEC-1999; 99WO-US28301.
 PR 02-DEC-1999; 99WO-US28365.
 PR 07-DEC-1999; 99US-0169495.
 PR 05-JAN-2000; 2000WO-US00219.
 PR 18-FEB-2000; 2000WO-US04341.
 PR 18-FEB-2000; 2000WO-US04342.
 PR 22-FEB-2000; 2000WO-US04414.
 PR 01-MAR-2000; 2000WO-US05601.
 PR 02-MAR-2000; 2000WO-US05841.
 PR 20-MAR-2000; 2000WO-US07377.
 PR 30-MAR-2000; 2000WO-US08439.
 PR 15-MAY-2000; 2000WO-US13358.
 PR 17-MAY-2000; 2000WO-US13705.

XX (GETH) GENENTECH INC.

XX Ashkenazi AJ, Baker KP, Botstein DA, Desnoyers LV, Eaton DL;
 PI Ferrara N, Fong S, Gao W, Gerber H, Gerltzen ME, Goddard A;
 PI Godowski PJ, Gunney AL, Kljavin IJ, Mather JP, Napier MA, Pan J;
 PI Paoni NF, Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM;
 PI Wood WI, Zhang Z;

XX WPI: 2001-050091/06.
 DR N-PSDB: AAC86968.

PT Isolated nucleic acid molecule encoding a PRO polypeptide which is a
 PT transmembrane polypeptide is useful for gene therapy and identification
 PT of related polypeptides -

PS Claim 12; Fig 10; 244pp; English.

XX The present sequence represents a human secreted and transmembrane
 CC polypeptide. The specification describes human polypeptides, designated
 CC PRO196, PRO444, PRO183, PRO185, PRO210, PRO215, PRO242, PRO288,
 CC PRO365, PRO1308, PRO1183, PRO1272, PRO1419, PRO4999, PRO7170,
 CC PRO248, PRO353, PRO1318, PRO1600, PRO9940, PRO533, PRO301, PRO187,
 CC PRO337, PRO4111, PRO4356, PRO246, PRO265, PRO941, PRO10096, PRO6003,
 CC PRO6004, PRO350, PRO2630 and PRO6309. The biological activity of cells

CC Hashimoto's thyroiditis, juvenile lymphocytic thyroiditis, atrophic
 CC thyroiditis), diabetes mellitus, immune-mediated renal disease
 CC (glomerulonephritis, tubulointerstitial nephritis), multiple sclerosis,
 CC idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome, chronic
 CC inflammatory demyelinating polyneuropathy, infectious hepatitis
 CC (hepatitis A, B, C, D, E and other non-hepatotropic viruses), autoimmune
 CC chronic active hepatitis, primary biliary cirrhosis, granulomatous
 CC hepatitis, and sclerosing cholangitis, inflammatory bowel disease
 CC (ulcerative colitis: Crohn's disease), gluten-sensitive enteropathy, and
 CC Whipple's disease. Autoimmune or immune-mediated skin diseases including,
 CC bullous skin diseases, erythema multiforme, contact dermatitis, psoriasis,
 CC asthma, allergic rhinitis, atopic dermatitis, food hypersensitivity,
 CC urticaria, eosinophilic pneumonia, idiopathic pulmonary fibrosis,
 CC hypersensitivity pneumonitis, and transplantation associated diseases
 CC (graft rejection, and graft-versus-host-disease). (1), its antagonists
 CC or fragment can also be used as an adjuvant in treatment of tumors.
 CC Antibodies against (1) can also be used for diagnosing such diseases.
 CC This sequence represents a human EGF-like homologue encoded by cDNA clone
 CC DNA32279 which is described in the invention.

Sequence 448 AA:

Query Match 99.88; Score 2527; DB 20; Length 448;
 Best Local Similarity 99.88; Pred. No. 8.1e-164;
 Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPEIKRLVTITIALCLPSRGNAGAOCTNGFDLRQSGQCLDIDECRTPEACRQDMCV 60
 DB 1 MPEIKRLVTITIALCLPSRGNAGAOCTNGFDLRQSGQCLDIDECRTPEACRQDMCV 60
 QY 61 NONGGYCIRTNTNVRGPRYSNPSTPSGPRPAAPLSPANPTISRLPCRGYOND 120
 DB 61 NONGGYCIRTNTNVRGPRYSNPSTPSGPRPAAPLSPANPTISRLPCRGYOND 120
 QY 121 ESNQCVVDECATSDHOCNPTQICINTEGTYCTSDGYWLEGGCLDIDECRYGCOOL 180
 DB 121 ESNQCVVDECATSDHOCNPTQICINTEGTYCTSDGYWLEGGCLDIDECRYGCOOL 180
 QY 121 ESNQCVVDECATSDHOCNPTQICINTEGTYCTSDGYWLEGGCLDIDECRYGCOOL 180
 DB 121 ESNQCVVDECATSDHOCNPTQICINTEGTYCTSDGYWLEGGCLDIDECRYGCOOL 180
 QY 181 CANVPGSYCTCNGPFTLNEDGRSCQDVNECATENPCVOTCVNTGSGFCRDPGELE 240
 DB 181 CANVPGSYCTCNGPFTLNEDGRSCQDVNECATENPCVOTCVNTGSGFCRDPGELE 240
 QY 241 DGVHCSMDDECSFSEFLCQHECVNPGTYFCSPGTYLLDDNRSCQDINECHRNHTCN 300
 DB 241 DGVHCSMDDECSFSEFLCQHECVNPGTYFCSPGTYLLDDNRSCQDINECHRNHTCN 300
 QY 301 LQOTCYNLOGGFKCIDPRCEPFLRISDNRCMCPAENPCRDQPTLLYRDMDVVSGRS 360
 DB 301 LQOTCYNLOGGFKCIDPRCEPFLRISDNRCMCPAENPCRDQPTLLYRDMDVVSGRS 360
 QY 361 VPADIFQWQATRRPGATYIRQIKSGNGREFYMGOTGPIATVTRPKGPREIOLDL 420
 DB 361 VPADIFQWQATRRPGATYIRQIKSGNGREFYMGOTGPIATVTRPKGPREIOLDL 420
 QY 421 EMITVNTVINFRGSSVIRLRIVSQYPF 448
 DB 421 EMITVNTVINFRGSSVIRLRIVSQYPF 448

RESULT 9
 AAU29227
 ID AAU29227 standard; Protein: 448 AA.

AAU29227;
 DT 18-DEC-2001 (first entry)

Human PRO polypeptide sequence #204.

XX PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
 KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
 KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
 KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.

XX Homo sapiens.
 OS
 XX WO200168848-A2.
 PN
 XX
 PD
 XX
 PF
 XX

20-SEP-2001.

28-FEB-2001; 2001MO-US06520.

PR 01-MAR-2000; 2000MO-US05601.
 PR 02-MAR-2000; 2000MO-US05841.
 PR 03-MAR-2000; 2000US-187202P.
 PR 06-MAR-2000; 2000US-186968P.
 PR 14-MAR-2000; 2000US-189320P.
 PR 14-MAR-2000; 2000US-189328P.
 PR 15-MAR-2000; 2000MO-US06884.
 PR 21-MAR-2000; 2000US-190828P.
 PR 21-MAR-2000; 2000US-191007P.
 PR 21-MAR-2000; 2000US-191048P.
 PR 21-MAR-2000; 2000US-191314P.
 PR 28-MAR-2000; 2000US-192655P.
 PR 29-MAR-2000; 2000US-193032P.
 PR 29-MAR-2000; 2000US-193053P.
 PR 30-MAR-2000; 2000MO-US08439.
 PR 04-APR-2000; 2000US-194449P.
 PR 04-APR-2000; 2000US-194647P.
 PR 11-APR-2000; 2000US-195975P.
 PR 11-APR-2000; 2000US-196008P.
 PR 11-APR-2000; 2000US-196187P.
 PR 11-APR-2000; 2000US-196690P.
 PR 11-APR-2000; 2000US-196820P.
 PR 18-APR-2000; 2000US-198121P.
 PR 18-APR-2000; 2000US-198585P.
 PR 25-APR-2000; 2000US-199397P.
 PR 25-APR-2000; 2000US-199550P.
 PR 25-APR-2000; 2000US-199634P.
 PR 03-MAY-2000; 2000US-201516P.
 PR 17-MAY-2000; 2000MO-US13705.
 PR 22-MAY-2000; 2000MO-US14042.
 PR 30-MAY-2000; 2000MO-US14941.
 PR 02-JUN-2000; 2000MO-US15264.
 PR 05-JUN-2000; 2000US-209832P.
 PR 28-JUL-2000; 2000MO-US20710.
 PR 22-AUG-2000; 2000US-0644848.
 PR 24-AUG-2000; 2000MO-US23328.
 PR 08-NOV-2000; 2000MO-US30952.
 PR 01-DEC-2000; 2000MO-US32678.
 PR 20-DEC-2000; 2000MO-US34956.

(GETH) GENENTECH INC.

Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
 Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;

WPI: 2001-602746/68.
 DR N-PSDB: AAS46128.

Novel nucleic acids encoding PRO polypeptides, used to diagnose the
 presence of tumours, such as prostate and breast tumours, in mammals and
 to screen for modulators of the compounds -

Claim 11; Fig 408; 774pp: English.

Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
 The PRO polypeptides and their associated nucleic acids can be used to
 detect the presence of a tumour in a mammal by comparing the level of
 expression of a PRO polypeptide in a test sample of cells from the animal
 and a control sample of normal cells, whereby a higher level of
 expression in the test sample indicates the presence of a tumour in the
 CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
 CC and rabbits but are preferably human. The polypeptides can be used to
 stimulate tumour necrosis factor (TNF) alpha release from human blood,
 CC when contacted with it. A specific polypeptide can be used to stimulate

PN US2001051358-A1.
 XX 13-DEC-2001.
 XX 25-MAR-1999; 99US-0275805.
 XX 11-APR-1997; 97US-0839525.
 PR 10-APR-1996; 96WO-US05247.
 XX (OLSEN/) OLSEN H S.
 PA (LIH/) LI H.
 PI Olsen HS, LI H;
 DR WPI; 2002-121417/16.
 DR N-PSDB; ABK13627.
 PT New nucleic acid encoding human extracellular/epidermal growth factor,
 PT useful for diagnosis and treatment of e.g. Marfan syndrome and wounds,
 PT also related polypeptides -
 XX
 XX Claim 10; Fig 1; 22pp; English.
 CC The invention relates to a novel polynucleotide which is at
 CC least 95% identical with a sequence (ATCC 97285) encoding mature human
 CC extracellular protein-like/epidermal growth factor (EGF)-like protein,
 CC EGF. Also included are the EGF EGF domains, a vector containing
 CC the polynucleotide, a host cell containing the vector, anti-EGF
 CC antibodies and antagonists of EGF. The polynucleotide is used for
 CC recombinant production of EGF, in gene therapy, as hybridisation probes,
 CC as antisense antagonists and for chromosome identification. The protein
 CC is used to treat patients who require EGF, to identify specific
 CC antagonists, used to treat conditions that require inhibition of EGF
 CC (e.g. vascular smooth muscle cell proliferation, Marfan syndrome, wound
 CC healing, neurological trauma, acquired immunodeficiency syndrome
 CC (AIDS)-related dementia, ocular disorders, kidney disorders, liver
 CC disorders, hair follicle growth promotion, burns, ulcers, corneal
 CC incisions, corneal inflammation, neoplasms and psoriasis), to raise
 CC specific antibodies and to characterise receptors. The present
 CC sequence represents EGF.
 CC
 SO Sequence 448 AA:
 Query Match 100.0%; Score 2533; DB 23; Length 448;
 Best Local Similarity 100.0%; Pred. No. 3.2e-164;
 Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 MPGKRLITVITIALCLPSPGNAOQCTNGFDLDRSGCCLDIDECRTPEACRGDMCV 60
 1 NONGGYLCIPRTNPNVYRGPNYSPYSGPYPAAPAPLSAPNYPITISPLICRFGYOMD 120
 61 NONGGYLCIPRTNPNVYRGPNYSPYSGPYPAAPAPLSAPNYPITISPLICRFGYOMD 120
 61 NONGGYLCIPRTNPNVYRGPNYSPYSGPYPAAPAPLSAPNYPITISPLICRFGYOMD 120
 121 ESNOCVAVDECATDSDHQCNPQICINTEGGYCTCTDGYMLLEGGCLDIDECRYGCOQL 180
 121 ESNOCVAVDECATDSDHQCNPQICINTEGGYCTCTDGYMLLEGGCLDIDECRYGCOQL 180
 121 ESNOCVAVDECATDSDHQCNPQICINTEGGYCTCTDGYMLLEGGCLDIDECRYGCOQL 180
 181 CANVPGSYSCNNGFTLNEDGRSCQDVNECATENPCVOTCVNTYSFICRCDPGYLEE 240
 181 CANVPGSYSCNNGFTLNEDGRSCQDVNECATENPCVOTCVNTYSFICRCDPGYLEE 240
 181 CANVPGSYSCNNGFTLNEDGRSCQDVNECATENPCVOTCVNTYSFICRCDPGYLEE 240
 241 DGVHCSMDDECSFESEFLCQHECVNPGTYFCSPGTYILLDNRSCQDINECHRNTCN 300
 241 DGVHCSMDDECSFESEFLCQHECVNPGTYFCSPGTYILLDNRSCQDINECHRNTCN 300
 241 DGVHCSMDDECSFESEFLCQHECVNPGTYFCSPGTYILLDNRSCQDINECHRNTCN 300
 301 LQGTCTYNLGGFKCIDPRICEEPYLRISDNRCMCPAENGCARDQPTILYRDMVYSGRS 360
 301 LQGTCTYNLGGFKCIDPRICEEPYLRISDNRCMCPAENGCARDQPTILYRDMVYSGRS 360
 301 LQGTCTYNLGGFKCIDPRICEEPYLRISDNRCMCPAENGCARDQPTILYRDMVYSGRS 360
 361 VPADIFQWQATRRPGAYIYIPIKSGNGREFYMQTGTISATVITRPIKGRREIQLD 420
 361 VPADIFQWQATRRPGAYIYIPIKSGNGREFYMQTGTISATVITRPIKGRREIQLD 420

DB 361 VPADIFQWQATRRPGAYIYIPIKSGNGREFYMQTGTISATVITRPIKGRREIQLD 420
 QY 421 EMITVNTVINRGSSVIRLRITYVSOYF 448
 DB 421 EMITVNTVINRGSSVIRLRITYVSOYF 448
 RESULT 8
 AAY08063
 ID AAY08063 standard; Protein: 448 AA.
 AC AAY08063;
 XX
 XX 11-SEP-2000 (first entry)
 DT
 DE Human EGF-like homologue protein (PRO217) encoded by DNA32279 cDNA.
 XX
 KW Inflammatory cell infiltration; immune response; T cell proliferation;
 KW anti-inflammatory; anti-autoimmune; anti-diabetic; spondyloarthritis;
 KW T cell-mediated disease; spondyloarthritis; sclerosis; renal disease;
 KW inflammatory myopathy; hemolytic anemia; thrombocytopenia; thyroiditis;
 KW diabetes mellitus; demyelinating polyneuropathy; Guillain-Barre syndrome;
 KW multiple sclerosis; polyneuropathy; hepatitis; cirrhosis; enteropathy;
 KW sclerosing cholangitis; inflammatory bowel disease; Whipple's disease;
 KW skin disease; dermatitis; psoriasis; asthma; allergic rhinitis; tumor;
 KW food hypersensitivity; urticaria; eosinophilic pneumonia; transplant;
 KW idiopathic pulmonary fibrosis; graft rejection; PRO245; human; PRO217;
 EGF-like.
 OS Homo sapiens.
 XX
 PN WO9914241-A2.
 XX 25-MAR-1999;
 PD 17-SEP-1998; 98WO-US19437.
 PE 17-SEP-1998; 98WO-US19437.
 XX 17-SEP-1997; 97US-0059119.
 XX 18-SEP-1997; 97US-0059263.
 PR 28-OCT-1997; 97US-0063550.
 PR 12-NOV-1997; 97US-0065186.
 PR 21-NOV-1997; 97US-0066364.
 PR 24-NOV-1997; 97US-0066770.
 PR 04-JUN-1998; 98US-0088026.
 PA (GETH) GENENTECH INC.
 PI Fong S, Goddard A, Gurney AL, Tumas D, Wood WI;
 DR WPI; 1999-229499/19.
 DR N-PSDB; AAX37670.
 PT Composition containing novel polypeptide PRO245, its agonist or
 PT antagonist -
 XX
 XX Example 1; Fig 6a; 177pp; English.
 PS
 XX This invention describes a novel composition containing (apart from a
 CC carrier or excipient), a novel PRO245 polypeptide (I), its agonist or
 CC antagonist, or their fragments, for modulating: (i) infiltration of
 CC inflammatory cells into tissue; (ii) an immune response; or (iii) T cell
 CC proliferation. The products of the invention have anti-inflammatory,
 CC (1)-(iii). The products of the invention have anti-inflammatory,
 CC anti-autoimmune and anti-diabetic activity. (I), and its (ant)agonists
 CC and their fragments, are used to treat immune-related diseases,
 CC particularly T cell-mediated diseases. The diseases treated include
 CC systemic lupus erythematosus, rheumatoid arthritis, juvenile chronic
 CC arthritis, spondyloarthritis, systemic sclerosis (scleroderma),
 CC idiopathic inflammatory myopathies (dermatomyositis, polymyositis),
 CC Sjogren's syndrome, systemic vasculitis, sarcoidosis, autoimmune
 CC hemolytic anemia (immune pancytopenia, paroxysmal nocturnal
 CC hemoglobinuria), autoimmune thrombocytopenia (idiopathic thrombocytopenic
 CC purpura immune-mediated thrombocytopenia), thyroiditis (Grave's disease,

PT smooth muscle cell growth, vasculogenesis, restenosis or
 PT atherosclerosis
 XX
 PS Claim 4: Page 46-47: 49pp; English.

CC This is the amino acid sequence of the human secreted protein AK647. The
 CC polynucleotide sequence was obtained from a human foetal kidney cDNA
 CC library. AK647 homologues in chicks and rodents are involved in aortic
 CC tissue development. The spatial and temporal distribution of AK647
 CC indicated that it acts as a modulator of smooth muscle cells in
 CC vasculogenesis during embryonic development. The primary structure of
 CC AK647 consists of multiple EGF domains. The AK647 protein can be used as
 CC a nutritional source or supplement. The protein shows both inhibitory and
 CC inducing, cytokine, cell proliferation and cell differentiation activity.
 CC The protein may also be used in the treatment of immune deficiencies and
 CC disorders, including severe combined immunodeficiency (SCID), HIV and
 CC other viral, bacterial and fungal infections. Regulation of immune
 CC responses may also be carried out by the AK647 protein. Other uses of the
 CC protein include a role in the regulation of haematopoiesis and in the
 CC treatment of myeloid and lymphoid cell deficiencies. Uses in bone,
 CC cartilage, tendon, ligament and nerve tissue regrowth are also possible,
 CC as well as for wound healing and in the treatment of ulcers and burns.
 CC The polynucleotides and proteins can be used for preventing, treating or
 CC ameliorating smooth muscle cell growth, vasculogenesis, restenosis,
 CC atherosclerosis, blood vessel remodelling and degeneration. The proteins
 CC may also have activin/inhibin, chemotactic/chemokinetic, haemostatic and
 CC thrombolytic, receptor/ligand, anti-inflammatory, cadherin/tumour
 CC invasion suppressor, and tumour inhibition activity. AK647 specific
 CC antibodies can be used for promoting smooth muscle cell growth or
 CC vasculogenesis. The proteins and polynucleotides can also be used for
 CC detection, diagnosis and drug screening.

SQ Sequence 448 AA;

Query Match 100.0%; Score 2533; DB 21; Length 448;

Best Local Similarity 100.0%; Pred. No. 3.2e-164; Mismatches 0; Indels 0; Gaps 0;

Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPGKIRLITVITLALCLPSPGNAQAQCTNGFDLDRSGGCLDIDERTIPEACRGDMCV 60
 DB 1 MPGKIRLITVITLALCLPSPGNAQAQCTNGFDLDRSGGCLDIDERTIPEACRGDMCV 60
 QY 61 NONGGYLCIRTPNRYRGPYSNPYSTPGYPAPAPLSAPNYPITSRPLICRFQYQMD 120
 DB 61 NONGGYLCIRTPNRYRGPYSNPYSTPGYPAPAPLSAPNYPITSRPLICRFQYQMD 120
 QY 121 ESNOCVDVDECATDSHOCNPTQICINTEGGYTCSTDGWMLLEGQCLDIDECRYGCOOL 180
 DB 121 ESNOCVDVDECATDSHOCNPTQICINTEGGYTCSTDGWMLLEGQCLDIDECRYGCOOL 180
 QY 181 CANVPGSYCTCNPFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGYELE 240
 DB 181 CANVPGSYCTCNPFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGYELE 240
 QY 241 DGVHCSMDSCSFSEFLCQHECVNPGTYFCSPGTYILLDNRSQDINECEHRNHTCN 300
 DB 241 DGVHCSMDSCSFSEFLCQHECVNPGTYFCSPGTYILLDNRSQDINECEHRNHTCN 300
 QY 301 LOOTCYMLGGGFKCIDPFRCEEPYLRISDNRCMPAENPGCRDOPFTIILYRDMVYSGRS 360
 DB 301 LOOTCYMLGGGFKCIDPFRCEEPYLRISDNRCMPAENPGCRDOPFTIILYRDMVYSGRS 360
 QY 361 VPADIFQMAATRRYPGAYYIFQIKSGNREFYMRQTPGISATLVTRPIKGRREIQLD 420
 DB 361 VPADIFQMAATRRYPGAYYIFQIKSGNREFYMRQTPGISATLVTRPIKGRREIQLD 420
 QY 421 EMITVNTVINFRGSSVIRLITVISOYR 448
 DB 421 EMITVNTVINFRGSSVIRLITVISOYR 448

RESULT 5
 AAY54989

ID AAY54989 standard; Protein; 448 AA.

AC AAY54989;

DT 15-FEB-2000 (first entry)

DE Full length human A55 protein sequence.

XX A55 protein; human; smooth muscle proliferation; tissue generation;
 KW vascular smooth muscle cell; arteriosclerosis; tissue repair; myoma;
 KW vascular endothelial thickening; haematopoietic cell-regulator; cytokine;
 KW percutaneous transluminal coronary angioplasty; blood coagulation; PRCA;
 KW actin; inhibin; chemotaxis; thrombosis; cadherin; therapy;
 KW tumour metastasis inhibitor.

OS Homo sapiens.

PN W09955864-A1.

PD 04-NOV-1999.

PF 28-APR-1999; 99WO-JP02284.

PR 28-APR-1998; 98JP-0119731.

PA (ONOV) ONO PHARM CO LTD.

PI Honjo T, Tashiro K, Nakamura T;

DR WPI: 2000-038647/03.

DR N-PSDB: AAZ40027.

PT Novel human polypeptides for treatment of, e.g. arteriosclerosis and
 myoma

PS Claim 1; Page 76-80; 87pp; Japanese.

CC This sequence is the human A55 protein of the invention. The protein
 CC can be used for the treatment of diseases due to abnormal proliferation
 CC of smooth muscle. The polypeptides can be used according their inhibition
 CC of the proliferation of vascular smooth muscle cells, particularly in
 CC treating arteriosclerosis or re-narrowing by vascular endothelial
 CC thickening after percutaneous transluminal coronary angioplasty (PTCA),
 CC or myoma, haematopoietic cell-regulatory activity, cytokine activity,
 CC tissue generation/repair activity, actin/inhibin activity, taxis
 CC and chemotaxis activity, blood coagulation/thrombotic activity,
 CC receptor/ligand activity, cadherin/tumour metastasis inhibiting activity;
 CC tumour inhibition, and as a nutrient.

SQ Sequence 448 AA;

Query Match 100.0%; Score 2533; DB 21; Length 448;

Best Local Similarity 100.0%; Pred. No. 3.2e-164; Mismatches 0; Indels 0; Gaps 0;

Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPGKIRLITVITLALCLPSPGNAQAQCTNGFDLDRSGGCLDIDERTIPEACRGDMCV 60
 DB 1 MPGKIRLITVITLALCLPSPGNAQAQCTNGFDLDRSGGCLDIDERTIPEACRGDMCV 60
 QY 61 NONGGYLCIRTPNRYRGPYSNPYSTPGYPAPAPLSAPNYPITSRPLICRFQYQMD 120
 DB 61 NONGGYLCIRTPNRYRGPYSNPYSTPGYPAPAPLSAPNYPITSRPLICRFQYQMD 120
 QY 121 ESNOCVDVDECATDSHOCNPTQICINTEGGYTCSTDGWMLLEGQCLDIDECRYGCOOL 180
 DB 121 ESNOCVDVDECATDSHOCNPTQICINTEGGYTCSTDGWMLLEGQCLDIDECRYGCOOL 180
 QY 181 CANVPGSYCTCNPFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGYELE 240
 DB 181 CANVPGSYCTCNPFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGYELE 240
 QY 241 DGVHCSMDSCSFSEFLCQHECVNPGTYFCSPGTYILLDNRSQDINECEHRNHTCN 300
 DB 241 DGVHCSMDSCSFSEFLCQHECVNPGTYFCSPGTYILLDNRSQDINECEHRNHTCN 300

QY 61 NONGGCLIPRTNPVRYGPNSTPYSGPYPAAPPLSAPNPTISRPLICRFQYMD 120
 Db 61 NONGGCLIPRTNPVRYGPNSTPYSGPYPAAPPLSAPNPTISRPLICRFQYMD 120
 QY 121 ESNOCVDECATDHSQCNPQTQICINTEGGYTCSCTDGYWLLBEGQCLDIDECRYGCOOL 180
 Db 121 ESNOCVDECATDHSQCNPQTQICINTEGGYTCSCTDGYWLLBEGQCLDIDECRYGCOOL 180
 QY 181 CANVGSYSCTCNPFTLNEDGRSCQDVNCAENPCVQTCVNTYGSFICRDPGEYLEE 240
 Db 181 CANVGSYSCTCNPFTLNEDGRSCQDVNCAENPCVQTCVNTYGSFICRDPGEYLEE 240
 QY 241 DGVHCSMDDECSFSEFLCQHECVNQPGTYFCSCPFGYILLDDNRSQDINCEHRNHTCN 300
 Db 241 DGVHCSMDDECSFSEFLCQHECVNQPGTYFCSCPFGYILLDDNRSQDINCEHRNHTCN 300
 QY 301 LQOTCYNLGGFKCIDPICEEPEYLRISDRNCMPAENPCGRDQPFITLYRDMQVVSGRS 360
 Db 301 LQOTCYNLGGFKCIDPICEEPEYLRISDRNCMPAENPCGRDQPFITLYRDMQVVSGRS 360
 QY 361 VPADIFQOMATTRYPGAYYIFQIKSGNEGREFYMRQTGPISATLVMTRPITKGPRIQDL 420
 Db 361 VPADIFQOMATTRYPGAYYIFQIKSGNEGREFYMRQTGPISATLVMTRPITKGPRIQDL 420
 QY 421 EMITVNTVINFRGSSVIRLRIYVSQYPF 448
 Db 421 EMITVNTVINFRGSSVIRLRIYVSQYPF 448
 RESULT 3
 AAM94281
 ID AAM94281 standard; Protein; 448 AA.
 AC AAM94281;
 DT 07-MAY-1999 (first entry)
 DE Human extracellular matrix protein (ECMP)-1.
 KW Extracellular matrix protein; ECMP-1; diagnosis; treatment; cancer;
 KM Immune disorder; human.
 OS Homo sapiens.
 PN MO9900410-A2.
 XX 07-JAN-1999.
 XX 23-JUN-1998; 98WO-US13012.
 XX 27-JUN-1997; 97US-0884072.
 XX (INCY-) INCYTE PHARM INC.
 XX Bandman O, Corley NC, Guegler KJ;
 XX WPI: 1999-095674/08.
 XX N-PSDB; AAX05359.
 XX New polynucleotide encoding extracellular matrix protein, ECMP-1 -
 PT useful in the diagnosis, prevention and treatment of immune
 XX disorders and cancer
 XX Claim 1; Fig 1A-G; 79pp; English.
 XX This represents a human extracellular matrix protein (ECMP)-1. Host
 CC cells containing a vector comprising the ECMP-1 nucleic acid are used
 CC for the recombinant production of the protein. ECMP-1 and its
 CC (ant)agonists, are useful in the diagnosis, prevention, and treatment
 CC of cancer and immune disorders.
 XX Sequence 448 AA:

Query Match 100.0%; Score 2533; DB 20; Length 448;
 Best Local Similarity 100.0%; Pred. No. 3.2e-164;
 Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGKIRILVTYIALCLPSPGNAOACTNGFDDLRQSGQCLDIDECRTIPACRGDMCV 60
 Db 1 MGKIRILVTYIALCLPSPGNAOACTNGFDDLRQSGQCLDIDECRTIPACRGDMCV 60
 QY 61 NONGGCLIPRTNPVRYGPNSTPYSGPYPAAPPLSAPNPTISRPLICRFQYMD 120
 Db 61 NONGGCLIPRTNPVRYGPNSTPYSGPYPAAPPLSAPNPTISRPLICRFQYMD 120
 QY 121 ESNOCVDECATDHSQCNPQTQICINTEGGYTCSCTDGYWLLBEGQCLDIDECRYGCOOL 180
 Db 121 ESNOCVDECATDHSQCNPQTQICINTEGGYTCSCTDGYWLLBEGQCLDIDECRYGCOOL 180
 QY 181 CANVGSYSCTCNPFTLNEDGRSCQDVNCAENPCVQTCVNTYGSFICRDPGEYLEE 240
 Db 181 CANVGSYSCTCNPFTLNEDGRSCQDVNCAENPCVQTCVNTYGSFICRDPGEYLEE 240
 QY 241 DGVHCSMDDECSFSEFLCQHECVNQPGTYFCSCPFGYILLDDNRSQDINCEHRNHTCN 300
 Db 241 DGVHCSMDDECSFSEFLCQHECVNQPGTYFCSCPFGYILLDDNRSQDINCEHRNHTCN 300
 QY 301 LQOTCYNLGGFKCIDPICEEPEYLRISDRNCMPAENPCGRDQPFITLYRDMQVVSGRS 360
 Db 301 LQOTCYNLGGFKCIDPICEEPEYLRISDRNCMPAENPCGRDQPFITLYRDMQVVSGRS 360
 QY 361 VPADIFQOMATTRYPGAYYIFQIKSGNEGREFYMRQTGPISATLVMTRPITKGPRIQDL 420
 Db 361 VPADIFQOMATTRYPGAYYIFQIKSGNEGREFYMRQTGPISATLVMTRPITKGPRIQDL 420
 QY 421 EMITVNTVINFRGSSVIRLRIYVSQYPF 448
 Db 421 EMITVNTVINFRGSSVIRLRIYVSQYPF 448
 RESULT 4
 AAY57058
 ID AAY57058 standard; Protein; 448 AA.
 AC AAY57058;
 DT 21-FEB-2000 (first entry)
 DE Amino acid sequence of the human secreted protein AK647.
 KW AK647; aortic tissue development; smooth muscle cell modulator; SCID;
 KW nutritional supplement; vasculogenesis; embryonic development; infection;
 KW cytokine activity; cell proliferation; cell differentiation; detect; HIV;
 KW immune deficiency; haematopoiesis regulation; tissue regrowth; diagnose;
 KW wound healing; restenosis; atherosclerosis; drug screen.
 OS Homo sapiens.
 PN MO9960125-A2.
 XX 25-NOV-1999.
 XX 18-MAY-1999; 99WO-US10931.
 XX 19-MAY-1998; 98US-0081002.
 XX 21-MAY-1998; 98US-0083002.
 XX (GEMV) GENETICS INST INC.
 XX Jacobs K, McCoy JM, Racle L, Lavallie E, Treacy M, Evans C;
 PI Agostino M, Lu Z, Werberg D;
 XX WPI: 2000-053298/04.
 XX N-PSDB; AAZ39892.
 XX Proteins, and their encoding polynucleotides, used for treating e.g.

FT Region /Label- EGF-4
 FT 272..314 /Label- EGF-5
 XX MO9846746-A1. (555)
 XX 22-OCT-1998.
 XX 11-APR-1997; 97MO-US06020.
 XX 11-APR-1997; 97MO-US06020.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX LI H, Olsen HS;
 XX WPI: 1998-568728/48.
 XX N-PSDB: AAV62432.
 DR New isolated extracellular/epidermal growth factor - used for
 PT regulating vascular smooth muscle cell proliferation, e.g. for
 DR enhancing neurological functions or treating neoplasia and other
 XX disorders.
 XX Claim 10a; Fig 1A-D; 62pp; English.

This sequence represents a novel human extracellular/epidermal growth factor-like protein, EGF. This protein can be used to regulate vascular smooth muscle cell proliferation and for restoration or enhancement of neurological functions diminished as a result of other damaging pathologies such as AIDS dementia. The protein can also be used to treat senile dementia, ocular disorders such as corneal inflammation, for targeting tumour cells, for treating kidney disorders, for liver regeneration or treating liver dysfunction, for treating wounds including all cutaneous wounds, corneal wounds, and injuries to the epithelial-lined hollow organs of the body or resulting from trauma such as burns, abrasions and cuts as well as from surgical procedures such as surgical incisions and skin grafting. The polypeptides can also be used for treating chronic conditions, such as chronic ulcers, diabetic ulcers, other non-healing (trophic) conditions, to treat Marfan syndrome, to promote hair follicular development, to stimulate growth and differentiation of various epidermal and epithelial cells in vivo and in vitro and to stimulate embryogenesis. Antagonists to EGF can be used to treat neoplasia such as cancers or tumours, skin disorders such as psoriasis or corneal inflammation. The products can also be used for identifying EGF receptors, detection, diagnosis and drug screening.

Sequence 448 AA:

Query Match 100.0%; Score 2533; DB 19; Length 448;
 Best Local Similarity 100.0%; Pred. No. 3.2e-164;
 Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPEKIRILVTYITIALCLPSPGNAOCTNGFDLDRSGQCCLDIDECRTIPEACRGDMKV 60
 DB 1 MPEKIRILVTYITIALCLPSPGNAOCTNGFDLDRSGQCCLDIDECRTIPEACRGDMKV 60
 QY 61 NONGGILCIPRTNPPVYRGPSNPSTPYSGPYPAAPPLSAPNPTIRPLICRGYOMD 120
 DB 61 NONGGILCIPRTNPPVYRGPSNPSTPYSGPYPAAPPLSAPNPTIRPLICRGYOMD 120
 QY 61 NONGGILCIPRTNPPVYRGPSNPSTPYSGPYPAAPPLSAPNPTIRPLICRGYOMD 120
 DB 61 NONGGILCIPRTNPPVYRGPSNPSTPYSGPYPAAPPLSAPNPTIRPLICRGYOMD 120
 QY 121 ESNQCVYVDECAADSHOQCNPTQICINTEGGYCTCTDGYWMLIEGQCLDIDECRYGCOOL 180
 DB 121 ESNQCVYVDECAADSHOQCNPTQICINTEGGYCTCTDGYWMLIEGQCLDIDECRYGCOOL 180
 QY 181 CANVPSSYSTCQNPGLTINEDGRSCQDVNECATENPCVQTCVNTYIGSTICRDPGELEE 240
 DB 181 CANVPSSYSTCQNPGLTINEDGRSCQDVNECATENPCVQTCVNTYIGSTICRDPGELEE 240
 QY 241 DGVHCSDMDECSSEFLQCHQECVNOQPGTYFCSCPPGYTLDDNRSCQDINCECHNHTCN 300
 DB 241 DGVHCSDMDECSSEFLQCHQECVNOQPGTYFCSCPPGYTLDDNRSCQDINCECHNHTCN 300

QY 301 LQOTCVNLQGGFKCIDPICEEPLRLISDNCKMCPAENPGCRDQPFLLYRDMVYSGRS 360
 DB 301 LQOTCVNLQGGFKCIDPICEEPLRLISDNCKMCPAENPGCRDQPFLLYRDMVYSGRS 360
 QY 361 VPADIFQMOATTRTPGAYYIFQIKSGNGBREFYMRQTPISATLVTRIKPREQLDL 420
 DB 361 VPADIFQMOATTRTPGAYYIFQIKSGNGBREFYMRQTPISATLVTRIKPREQLDL 420
 QY 421 EMITVNTVINFGSSVTRLRITYVSQYPF 448
 DB 421 EMITVNTVINFGSSVTRLRITYVSQYPF 448

RESULT 2

AAW95709
 ID AAW95709 standard; protein: 448 AA.

AAW95709;

21-JUN-1999 (first entry)

Homo sapiens fetal kidney clone AK647 secreted protein.

Secreted protein: fetal kidney.

Homo sapiens.

MO9900405-A1. (555)

07-JAN-1999.

29-JUN-1998; 98MO-US13530.

30-JUN-1997; 97US-0885610.

(GEMT) GENETICS INST INC.

Agostino MJ, Evans C, Jacobs K, Lavallie ER, McCoy JM;
 PI Merberg D, Racle LA, Treacy M;

WPI: 1998-095671/08.
 N-PSDB: AAX07567.

New polynucleotides encoding secreted human proteins - are derived from foetal kidney or adult retina cDNA libraries, used as, e.g. potential vaccines

Claim 11; Pages 52-54; 76pp; English.

The sequence is that of a secreted protein from a human fetal kidney clone AK296. Such a sequence is predicted to have biological activities which would make them suitable for treating, preventing or ameliorating medical conditions in humans and animals, although no supporting data is given. Suggested activities include nutritional activity, cytokine and cell proliferation/differentiation activity, immune stimulating (e.g. as vaccines) or suppressing activity, haematopoiesis regulating activity, tissue growth activity, activin/inhibin regulating activity, chemotactic/chemokinetic activity, and thrombolytic activity, receptor/ligand activity, anti-inflammatory activity, cadherin/tumour invasion suppressor activity, and tumour inhibition activity. It is also stated to be useful for gene therapy.

Sequence 448 AA:

Query Match 100.0%; Score 2533; DB 20; Length 448;
 Best Local Similarity 100.0%; Pred. No. 3.2e-164;
 Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPEKIRILVTYITIALCLPSPGNAOCTNGFDLDRSGQCCLDIDECRTIPEACRGDMKV 60
 DB 1 MPEKIRILVTYITIALCLPSPGNAOCTNGFDLDRSGQCCLDIDECRTIPEACRGDMKV 60

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 11:59:37 ; Search time 32.4041 Seconds
(without alignments)
1842.243 Million cell updates/sec

Title: US-09-674-379A-13
Perfect score: 2533
Sequence: 1 MEPIKRLVTITLALCLPSP.....INFGSSVIRLRITYSQYRF 448

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Archived: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*

1:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2533	100.0	448	19	AAW79739 Human EGF protein
2	2533	100.0	448	20	AAW95709 Homo sapiens fetal
3	2533	100.0	448	20	AAW94281 Homo extracellular
4	2533	100.0	448	21	AAV57058 Amino acid sequenc
5	2533	100.0	448	21	AAV54989 Full length human
6	2533	100.0	448	22	AAW93573 Human polypeptide,
7	2533	100.0	448	23	AAU75494 Human extracellular
8	2527	99.8	448	20	AAV08063 Human EGF-like hom
9	2527	99.8	448	22	AAU29227 Human PRO polypept
10	2527	99.8	448	22	AAW31183 Amino acid sequenc

11	2405	94.9	448	21	AAV56750 Smooth muscle prol
12	2405	94.9	448	21	AAV54990 Full length mouse
13	2376	93.8	461	21	AAV56752 Smooth muscle prol
14	2376	93.8	461	21	AAV54991 Full length mouse
15	2302	90.9	423	21	AAV56751 Smooth muscle prol
16	2302	90.9	423	21	AAV56753 Smooth muscle prol
17	2230	88.0	392	18	AAW31705 Human extracellular
18	1827	72.1	335	21	AAV76008 Rat EGF extracellu
19	1827	72.1	335	22	AAV55947 Skin cell protein,
20	1827	72.1	335	22	AAV72147 Rat protein isolat
21	1289	50.9	443	18	AAW32110 Human extracellular
22	1289	50.9	443	20	AAV16587 Extracellular prot
23	1289	50.9	443	21	AAV33418 Human PRO226 prote
24	1289	50.9	443	21	AAV84707 A human p53 mutant
25	1289	50.9	443	21	AAV55850 Human PRO226 polyp
26	1289	50.9	443	22	AAU12330 Human PRO226 polyp
27	1289	50.9	443	23	AAU086130 Human protein sequ
28	1283	50.7	443	22	AAV84706 Amino acid sequenc
29	1282	50.6	443	21	AAV84706 Lung cancer associ
30	1237.5	48.9	433	21	AAV58353 Human polypeptide,
31	1106	43.7	201	22	AAW93220 Human EFEMP1, Hom
32	1101	43.5	493	22	AAV72892 Human extracellular
33	1101	43.5	493	22	AAV84707 The C-terminal of
34	992.5	39.2	295	21	AAV84705 C-terminal of p53
35	989.5	39.1	295	21	AAV84708 Human EGF-like pro
36	963.5	38.0	350	20	AAV08066 Extracellular prot
37	963.5	38.0	387	22	AAV68188 Human EGF extracel
38	912.5	36.0	274	21	AAV76081 Skin cell protein,
39	912.5	36.0	274	22	AAV56020 Human protein Isol
40	912.5	36.0	274	23	AAV72220 Secreted protein e
41	819	32.3	413	22	AAV80391 Human albumin fusi
42	819	32.3	413	23	AAV65300 Gene #21 associate
43	819	32.3	451	22	AAV80440 Fibrillin C, Homo s
44	737	29.1	683	12	AAW11150 Human fibrillin type
45	737	29.1	683	18	AAW27600

ALIGNMENTS

RESULT 1	
ID	AAW79739
XX	AAW79739 standard; Protein; 448 AA.
XX	AAW79739;
XX	25-JAN-1999 (first entry)
DE	Human EGF protein.
XX	Extracellular/epidermal growth factor-like protein; EGF; human; liver;
KW	vascular smooth muscle cell proliferation; neurology; pathology; AIDS;
KW	dementia; ocular; disorder; cornea; inflammation; tumor cell; kidney;
KW	wound; epithelium; surgery; skin graft; ulcer; diabetic; trophic;
KW	Marfan syndrome; treatment; hair follicle; embryogenesis; neoplasia;
KW	epidermal cell; cancer; psoriasis; detection.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Peptide
FT	Protein
FT	Location/Qualifiers
FT	1..25
FT	/label= signal
FT	26..448
FT	/label= EGF
FT	/note= "Extracellular/epidermal growth factor-like protein"
FT	
FT	Region
FT	112..153
FT	/label= EGF-1
FT	154..190
FT	/label= EGF-2
FT	191..230
FT	/label= EGF-3
FT	231..271
FT	Region

```
QY 1 MEGIKRILVYTIILALCLPSPGNAQAQCTNGFDLDRSGQCLDIDECRTTPEACRGDMCV 60
Db 1 MEGIKRILVYTIILALCLPSPGNAQAQCTNGFDLDRSGQCLDIDECRTTPEACRGDMCV 60
QY 61 NONGGYLICIRTPNVPYRGPSNPNSTPYSGPYPAAPPLSAPNYPPTISRPLICRGYOMD 120
Db 61 NONGGYLICIRTPNVPYRGPSNPNSTPYSGPYPAAPPLSAPNYPPTISRPLICRGYOMD 120
QY 121 ESNQCVNDECATDSDHOCNPTQICINTEGYYTCSCTDGYWLLGQCLDIDECRYGCOOL 180
Db 121 ESNQCVNDECATDSDHOCNPTQICINTEGYYTCSCTDGYWLLGQCLDIDECRYGCOOL 180
QY 181 CANVPGSYSCTCNPGFTLNEDGRSCQDVNECATENPCVQTCVNTYGSFICRCDPGYELE 240
Db 181 CANVPGSYSCTCNPGFTLNEDGRSCQDVNECATENPCVQTCVNTYGSFICRCDPGYELE 240
QY 241 DGVHCSMDMECSFSEFLCQHECVNQPGTYFCSCPFGYLLDDNRSCODINECEHRNHTCN 300
Db 241 DGVHCSMDMECSFSEFLCQHECVNQPGTYFCSCPFGYLLDDNRSCODINECEHRNHTCN 300
QY 301 LOOTCYNLOGGFKCIDPIRCEEPYLRISDNRCMCPAENPGCRDOPFTILYRDMDVYSGRS 360
Db 301 LOOTCYNLOGGFKCIDPIRCEEPYLRISDNRCMCPAENPGCRDOPFTILYRDMDVYSGRS 360
QY 361 VPADIFOMQATTRYPGAYYIFQIKSGNEGREFYMRQGTGPISATLVMTRPPIKGPRIQIDL 420
Db 361 VPADIFOMQATTRYPGAYYIFQIKSGNEGREFYMRQGTGPISATLVMTRPPIKGPRIQIDL 420
QY 421 EMITVNTVINFRGSSVYRLRIYVSQYPF 448
Db 421 EMITVNTVINFRGSSVYRLRIYVSQYPF 448
```

Search completed: July 3, 2003, 18:26:16
Job time : 89.4684 secs

;; CURRENT APPLICATION NUMBER: US/10/176,757
;; CURRENT FILING DATE: 2002-06-20
;; Prior Application removed - See file Wrapper or Palm
;; NUMBER OF SEQ ID NOS: 612
;; SEQ ID NO 408
;; LENGTH: 448
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-176-757-408

Query Match 99.8%; Score 2527; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2.1e-164;
Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGKIRILVTITALLCPSPGNAOACTNGFDLROSGQCLDIDECRTIPEACRGDMCV 60
DB 1 MGKIRILVTITALLCPSPGNAOACTNGFDLROSGQCLDIDECRTIPEACRGDMCV 60
Y 61 NONGGYLCIPRTNPNVYRGPNYSPYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
DB 61 NONGGYLCIPRTNPNVYRGPNYSPYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
QY 121 ESNQCVYVDECATDSDHOCNPTQICINTEGGYTCSTDGWYLLGQCCLDIDECRYGCOOL 180
DB 121 ESNQCVYVDECATDSDHOCNPTQICINTEGGYTCSTDGWYLLGQCCLDIDECRYGCOOL 180
QY 181 CANVPGSYCTCNPFTLNEDGRSCODVNECATENPCVQTCVNTYGSFICRCDPGYELEE 240
DB 181 CANVPGSYCTCNPFTLNEDGRSCODVNECATENPCVQTCVNTYGSFICRCDPGYELEE 240
QY 241 DGVHCSMDCECSFSEFLQHECVNQPCTYFCSCPBGYILLDNRSCQDINECEHRNHTCN 300
DB 241 DGVHCSMDCECSFSEFLQHECVNQPCTYFCSCPBGYILLDNRSCQDINECEHRNHTCN 300
QY 301 LOOTCYNLOGGFKCIDIRCEEPYLRISDNRCMCAENPGCDOPFTILYRMDVYSGRS 360
DB 301 LOOTCYNLOGGFKCIDIRCEEPYLRISDNRCMCAENPGCDOPFTILYRMDVYSGRS 360
QY 361 VPADIFOMQATTRPGAYYIFQIKSGNEGREFYMRQTPISATLVMPRIKGPRIQIDL 420
DB 361 VPADIFOMQATTRPGAYYIFQIKSGNEGREFYMRQTPISATLVMPRIKGPRIQIDL 420
QY 421 EMITVNTVINFRGSSVIRLRIVSOYPF 448
DB 421 EMITVNTVINFRGSSVIRLRIVSOYPF 448

RESULT 14

US-10-176-913-408
;; Sequence 408, Application US/10176913
;; Publication No. US20030022298A1
;; GENERAL INFORMATION:
;; APPLICANT: Baker, Kevin P.
;; APPLICANT: Chen, Jian
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Pan, James
;; APPLICANT: Smith, Victoria
;; APPLICANT: Watanabe, Colin K.
;; APPLICANT: Wood, William I.
;; APPLICANT: Zhang, Zemin
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P3430R1C66
;; CURRENT APPLICATION NUMBER: US/10/176,913
;; CURRENT FILING DATE: 2002-06-20
;; Prior Application removed - See file Wrapper or Palm
;; NUMBER OF SEQ ID NOS: 612
;; SEQ ID NO 408
;; LENGTH: 448
;; TYPE: PRT

;; ORGANISM: Homo Sapien
US-10-176-913-408

Query Match 99.8%; Score 2527; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2.1e-164;
Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGKIRILVTITALLCPSPGNAOACTNGFDLROSGQCLDIDECRTIPEACRGDMCV 60
DB 1 MGKIRILVTITALLCPSPGNAOACTNGFDLROSGQCLDIDECRTIPEACRGDMCV 60
QY 61 NONGGYLCIPRTNPNVYRGPNYSPYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
DB 61 NONGGYLCIPRTNPNVYRGPNYSPYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
QY 121 ESNQCVYVDECATDSDHOCNPTQICINTEGGYTCSTDGWYLLGQCCLDIDECRYGCOOL 180
DB 121 ESNQCVYVDECATDSDHOCNPTQICINTEGGYTCSTDGWYLLGQCCLDIDECRYGCOOL 180
QY 181 CANVPGSYCTCNPFTLNEDGRSCODVNECATENPCVQTCVNTYGSFICRCDPGYELEE 240
DB 181 CANVPGSYCTCNPFTLNEDGRSCODVNECATENPCVQTCVNTYGSFICRCDPGYELEE 240
QY 241 DGVHCSMDCECSFSEFLQHECVNQPCTYFCSCPBGYILLDNRSCQDINECEHRNHTCN 300
DB 241 DGVHCSMDCECSFSEFLQHECVNQPCTYFCSCPBGYILLDNRSCQDINECEHRNHTCN 300
QY 301 LOOTCYNLOGGFKCIDIRCEEPYLRISDNRCMCAENPGCDOPFTILYRMDVYSGRS 360
DB 301 LOOTCYNLOGGFKCIDIRCEEPYLRISDNRCMCAENPGCDOPFTILYRMDVYSGRS 360
QY 361 VPADIFOMQATTRPGAYYIFQIKSGNEGREFYMRQTPISATLVMPRIKGPRIQIDL 420
DB 361 VPADIFOMQATTRPGAYYIFQIKSGNEGREFYMRQTPISATLVMPRIKGPRIQIDL 420
QY 421 EMITVNTVINFRGSSVIRLRIVSOYPF 448
DB 421 EMITVNTVINFRGSSVIRLRIVSOYPF 448

RESULT 15

US-10-180-552-408
;; Sequence 408, Application US/10180552
;; Publication No. US20030022300A1
;; GENERAL INFORMATION:
;; APPLICANT: Baker, Kevin P.
;; APPLICANT: Chen, Jian
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Pan, James
;; APPLICANT: Smith, Victoria
;; APPLICANT: Watanabe, Colin K.
;; APPLICANT: Wood, William I.
;; APPLICANT: Zhang, Zemin
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P3430R1C153
;; CURRENT APPLICATION NUMBER: US/10/180,552
;; CURRENT FILING DATE: 2002-06-25
;; Prior Application removed - See file Wrapper or Palm
;; NUMBER OF SEQ ID NOS: 612
;; SEQ ID NO 408
;; LENGTH: 448
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-180-552-408

Query Match 99.8%; Score 2527; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2.1e-164;
Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C60
; CURRENT APPLICATION NUMBER: US/10/175,752
; PRIOR APPLICATION: 2002-06-19
; NUMBER OF SEQ ID NOS: 612
; LENGTH: 448
; TYPE: PRP
; ORGANISM: Homo Sapien
; US-10-175-752-408

```

```

Query Match      99.8%; Score 2527; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2.1e-164;
Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 MPGIRILVTYITIALCLPSPGNAOACTNGFDLDROSGCCLDIDECRTIPEACRGDMCV 60
DB 1 MPGIRILVTYITIALCLPSPGNAOACTNGFDLDROSGCCLDIDECRTIPEACRGDMCV 60
QY 61 NONGGYLCIPRTNPVYRGPSNPSTYSGPYPAAPLAPNPTISRLICRGYQMD 120
DB 61 NONGGYLCIPRTNPVYRGPSNPSTYSGPYPAAPLAPNPTISRLICRGYQMD 120
QY 121 ESNOCVYDDECATSHOCPNPQICINTEGGYTCSTDYWLLEGCLDIDECRYGCOOL 180
DB 121 ESNOCVYDDECATSHOCPNPQICINTEGGYTCSTDYWLLEGCLDIDECRYGCOOL 180
QY 121 ESNOCVYDDECATSHOCPNPQICINTEGGYTCSTDYWLLEGCLDIDECRYGCOOL 180
DB 121 ESNOCVYDDECATSHOCPNPQICINTEGGYTCSTDYWLLEGCLDIDECRYGCOOL 180
QY 181 CANVPGSYCTCNGFTLNEDGRSCQDVNECATENPCVQTCVNTYGSFLICRCDPGYELEE 240
DB 181 CANVPGSYCTCNGFTLNEDGRSCQDVNECATENPCVQTCVNTYGSFLICRCDPGYELEE 240
QY 241 DGVCSDMDCESEFELQHCVCVNOPTGYFCSCPFGYILLDDNRSCQDINECEHNRHTCN 300
DB 241 DGVCSDMDCESEFELQHCVCVNOPTGYFCSCPFGYILLDDNRSCQDINECEHNRHTCN 300
QY 301 LOOTCYNLGGFKCIDPICEEPYLRISDNRCMCPAENPGCRDQFTILYRMDVYSGRS 360
DB 301 LOOTCYNLGGFKCIDPICEEPYLRISDNRCMCPAENPGCRDQFTILYRMDVYSGRS 360
QY 361 VPADIFOMQATTRYPGAYYIFQIKSGNEGREFYMRQTPISATLVMTPIKGPRETQDL 420
DB 361 VPADIFOMQATTRYPGAYYIFQIKSGNEGREFYMRQTPISATLVMTPIKGPRETQDL 420
QY 421 EMITVNTVINFRGSSVIRLRIVSOYRF 448
DB 421 EMITVNTVINFRGSSVIRLRIVSOYRF 448

```

```

RESULT 12
; US-10-176-482-408
; Sequence 408, Application US/10176482
; Publication No. US20030022296A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James

```

```

; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C70
; CURRENT APPLICATION NUMBER: US/10/176,482
; PRIOR APPLICATION: 2002-06-20
; NUMBER OF SEQ ID NOS: 612
; LENGTH: 448
; TYPE: PRP
; ORGANISM: Homo Sapien
; US-10-176-482-408

```

```

Query Match      99.8%; Score 2527; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2.1e-164;
Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 MPGIRILVTYITIALCLPSPGNAOACTNGFDLDROSGCCLDIDECRTIPEACRGDMCV 60
DB 1 MPGIRILVTYITIALCLPSPGNAOACTNGFDLDROSGCCLDIDECRTIPEACRGDMCV 60
QY 61 NONGGYLCIPRTNPVYRGPSNPSTYSGPYPAAPLAPNPTISRLICRGYQMD 120
DB 61 NONGGYLCIPRTNPVYRGPSNPSTYSGPYPAAPLAPNPTISRLICRGYQMD 120
QY 121 ESNOCVYDDECATSHOCPNPQICINTEGGYTCSTDYWLLEGCLDIDECRYGCOOL 180
DB 121 ESNOCVYDDECATSHOCPNPQICINTEGGYTCSTDYWLLEGCLDIDECRYGCOOL 180
QY 181 CANVPGSYCTCNGFTLNEDGRSCQDVNECATENPCVQTCVNTYGSFLICRCDPGYELEE 240
DB 181 CANVPGSYCTCNGFTLNEDGRSCQDVNECATENPCVQTCVNTYGSFLICRCDPGYELEE 240
QY 241 DGVCSDMDCESEFELQHCVCVNOPTGYFCSCPFGYILLDDNRSCQDINECEHNRHTCN 300
DB 241 DGVCSDMDCESEFELQHCVCVNOPTGYFCSCPFGYILLDDNRSCQDINECEHNRHTCN 300
QY 301 LOOTCYNLGGFKCIDPICEEPYLRISDNRCMCPAENPGCRDQFTILYRMDVYSGRS 360
DB 301 LOOTCYNLGGFKCIDPICEEPYLRISDNRCMCPAENPGCRDQFTILYRMDVYSGRS 360
QY 361 VPADIFOMQATTRYPGAYYIFQIKSGNEGREFYMRQTPISATLVMTPIKGPRETQDL 420
DB 361 VPADIFOMQATTRYPGAYYIFQIKSGNEGREFYMRQTPISATLVMTPIKGPRETQDL 420
QY 421 EMITVNTVINFRGSSVIRLRIVSOYRF 448
DB 421 EMITVNTVINFRGSSVIRLRIVSOYRF 448

```

```

RESULT 13
; US-10-176-757-408
; Sequence 408, Application US/10176757
; Publication No. US20030022297A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C86

```

QY 361 VPADIFOMQATRRYPGAYYIFQIKSGNEGREFYMQTGPISATLVWTRPIKGPRTIQLDL 420
Db 361 VPADIFOMQATRRYPGAYYIFQIKSGNEGREFYMQTGPISATLVWTRPIKGPRTIQLDL 420
QY 421 EMITVNTVINRGSSVIRLRIVYSQYPF 448
Db 421 EMITVNTVINRGSSVIRLRIVYSQYPF 448

RESULT 9

US-10-173-706-408
; Sequence 408, Application US/10173706
; Publication No. US20030022293A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C7
; CURRENT APPLICATION NUMBER: US/10/173,706
; PRIOR FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 408
; LENGTH: 448
; TYPE: PRF
; ORGANISM: Homo Sapien
US-10-173-706-408

Query Match 99.8%; Score 2527; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2,1e-164;

Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGKIRILVTITLALCLPSPGNAOACTNGFDDROSGCGLDIDECRTIPEACRGDMVCV 60
Db 1 MGKIRILVTITLALCLPSPGNAOACTNGFDDROSGCGLDIDECRTIPEACRGDMVCV 60
QY 61 NONGGYLCIPRTNPYRGPSNPNSTYSGPYPAAPPLSAPNPTISRLICRFQYMD 120
Db 61 NONGGYLCIPRTNPYRGPSNPNSTYSGPYPAAPPLSAPNPTISRLICRFQYMD 120
QY 121 ESNOCVDVDECATDHSQCNPQTICINTEGGYTCSCTDGYWLEGGCGLDIDECRYGYCOQL 180
Db 121 ESNOCVDVDECATDHSQCNPQTICINTEGGYTCSCTDGYWLEGGCGLDIDECRYGYCOQL 180
QY 181 CANVGSYSCTCNPFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGYLEE 240
Db 181 CANVGSYSCTCNPFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGYLEE 240
QY 241 DGVHCSMDDECSFSEFLCOHECVNPGTYFCSCPFGYILLDDNRSCODINECHRNHTCN 300
Db 241 DGVHCSMDDECSFSEFLCOHECVNPGTYFCSCPFGYILLDDNRSCODINECHRNHTCN 300
QY 301 LGQTCYNLQGGFKCIDPIRCEEPYLRISDNRCMCPAENPGCRDQPTIILYRDMDVYSGRS 360
Db 301 LGQTCYNLQGGFKCIDPIRCEEPYLRISDNRCMCPAENPGCRDQPTIILYRDMDVYSGRS 360
QY 361 VPADIFOMQATRRYPGAYYIFQIKSGNEGREFYMQTGPISATLVWTRPIKGPRTIQLDL 420
Db 361 VPADIFOMQATRRYPGAYYIFQIKSGNEGREFYMQTGPISATLVWTRPIKGPRTIQLDL 420
QY 421 EMITVNTVINRGSSVIRLRIVYSQYPF 448
Db 421 EMITVNTVINRGSSVIRLRIVYSQYPF 448

RESULT 10

US-10-175-738-408
; Sequence 408, Application US/10175738
; Publication No. US20030022294A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C45
; CURRENT APPLICATION NUMBER: US/10/175,738
; PRIOR FILING DATE: 2002-06-19
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 408
; LENGTH: 448
; TYPE: PRF
; ORGANISM: Homo Sapien
US-10-175-738-408

Query Match 99.8%; Score 2527; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2,1e-164;

Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGKIRILVTITLALCLPSPGNAOACTNGFDDROSGCGLDIDECRTIPEACRGDMVCV 60
Db 1 MGKIRILVTITLALCLPSPGNAOACTNGFDDROSGCGLDIDECRTIPEACRGDMVCV 60
QY 61 NONGGYLCIPRTNPYRGPSNPNSTYSGPYPAAPPLSAPNPTISRLICRFQYMD 120
Db 61 NONGGYLCIPRTNPYRGPSNPNSTYSGPYPAAPPLSAPNPTISRLICRFQYMD 120
QY 121 ESNOCVDVDECATDHSQCNPQTICINTEGGYTCSCTDGYWLEGGCGLDIDECRYGYCOQL 180
Db 121 ESNOCVDVDECATDHSQCNPQTICINTEGGYTCSCTDGYWLEGGCGLDIDECRYGYCOQL 180
QY 181 CANVGSYSCTCNPFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGYLEE 240
Db 181 CANVGSYSCTCNPFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGYLEE 240
QY 241 DGVHCSMDDECSFSEFLCOHECVNPGTYFCSCPFGYILLDDNRSCODINECHRNHTCN 300
Db 241 DGVHCSMDDECSFSEFLCOHECVNPGTYFCSCPFGYILLDDNRSCODINECHRNHTCN 300
QY 301 LGQTCYNLQGGFKCIDPIRCEEPYLRISDNRCMCPAENPGCRDQPTIILYRDMDVYSGRS 360
Db 301 LGQTCYNLQGGFKCIDPIRCEEPYLRISDNRCMCPAENPGCRDQPTIILYRDMDVYSGRS 360
QY 361 VPADIFOMQATRRYPGAYYIFQIKSGNEGREFYMQTGPISATLVWTRPIKGPRTIQLDL 420
Db 361 VPADIFOMQATRRYPGAYYIFQIKSGNEGREFYMQTGPISATLVWTRPIKGPRTIQLDL 420
QY 421 EMITVNTVINRGSSVIRLRIVYSQYPF 448
Db 421 EMITVNTVINRGSSVIRLRIVYSQYPF 448

RESULT 11

US-10-175-752-408
; Sequence 408, Application US/10175752
; Publication No. US20030022295A1
; GENERAL INFORMATION:

Db 121 ESNQCVDBECATDSHOCNPTQICINTEGGYTCSTGDTGWLLEGGCLDIDECRYGCAQL 180
QY 181 CANVPGSYSTCTCNGPFTLNEDGRSCQDVNECATENPCVQTCVNTYGSFICRCDPGELEE 240
Db 181 CANVPGSYSTCTCNGPFTLNEDGRSCQDVNECATENPCVQTCVNTYGSFICRCDPGELEE 240
QY 241 DGVHCSMDMECSSEFLCOHECVNQPGTYFCSCPFGYILLDDNRSCQDINECEHNRHNCN 300
Db 241 DGVHCSMDMECSSEFLCOHECVNQPGTYFCSCPFGYILLDDNRSCQDINECEHNRHNCN 300
QY 301 LOOTCYNLOGGFRCIDPICEEPLRISDNRCMCPAENPGCRDOPFTILYRDMDVSGRS 360
Db 301 LOOTCYNLOGGFRCIDPICEEPLRISDNRCMCPAENPGCRDOPFTILYRDMDVSGRS 360
QY 361 VPADIFOMATTYRPGAYYIFQIKSGNEGREFYRQTPISATLVMPRIKGPRIQDL 420
Db 361 VPADIFOMATTYRPGAYYIFQIKSGNEGREFYRQTPISATLVMPRIKGPRIQDL 420
QY 421 EMITVTVINFRGSSVIRLRIRIYVSQYDF 448
Db 421 EMITVTVINFRGSSVIRLRIRIYVSQYDF 448

RESULT 7

US-10-176-758-408
; Sequence 408, Application US/10176758
; Publication No. US20030008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Matanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C104
; CURRENT APPLICATION NUMBER: US/10/176,758
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 408
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-758-408

Query Match 99.8%; Score 2527; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2.1e-164;
Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPGIKRLITVITIALCLPSPGNAOACTNGFDLDRSGCCLDIDECRTIPEACRGDMCV 60
Db 1 MPGIKRLITVITIALCLPSPGNAOACTNGFDLDRSGCCLDIDECRTIPEACRGDMCV 60
QY 61 NONGGYLCIPRTNPNVYRGPSNPSTYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
Db 61 NONGGYLCIPRTNPNVYRGPSNPSTYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
QY 121 ESNQCVDBECATDSHOCNPTQICINTEGGYTCSTGDTGWLLEGGCLDIDECRYGCAQL 180
Db 121 ESNQCVDBECATDSHOCNPTQICINTEGGYTCSTGDTGWLLEGGCLDIDECRYGCAQL 180
QY 181 CANVPGSYSTCTCNGPFTLNEDGRSCQDVNECATENPCVQTCVNTYGSFICRCDPGELEE 240
Db 181 CANVPGSYSTCTCNGPFTLNEDGRSCQDVNECATENPCVQTCVNTYGSFICRCDPGELEE 240
QY 241 DGVHCSMDMECSSEFLCOHECVNQPGTYFCSCPFGYILLDDNRSCQDINECEHNRHNCN 300
Db 241 DGVHCSMDMECSSEFLCOHECVNQPGTYFCSCPFGYILLDDNRSCQDINECEHNRHNCN 300

Db 241 DGVHCSMDMECSSEFLCOHECVNQPGTYFCSCPFGYILLDDNRSCQDINECEHNRHNCN 300
QY 301 LOOTCYNLOGGFRCIDPICEEPLRISDNRCMCPAENPGCRDOPFTILYRDMDVSGRS 360
Db 301 LOOTCYNLOGGFRCIDPICEEPLRISDNRCMCPAENPGCRDOPFTILYRDMDVSGRS 360
QY 361 VPADIFOMATTYRPGAYYIFQIKSGNEGREFYRQTPISATLVMPRIKGPRIQDL 420
Db 361 VPADIFOMATTYRPGAYYIFQIKSGNEGREFYRQTPISATLVMPRIKGPRIQDL 420
QY 421 EMITVTVINFRGSSVIRLRIRIYVSQYDF 448
Db 421 EMITVTVINFRGSSVIRLRIRIYVSQYDF 448

RESULT 8

US-10-175-737-408
; Sequence 408, Application US/10175737
; Publication No. US20030013153A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Matanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C150
; CURRENT APPLICATION NUMBER: US/10/175,737
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 408
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-737-408

Query Match 99.8%; Score 2527; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2.1e-164;
Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPGIKRLITVITIALCLPSPGNAOACTNGFDLDRSGCCLDIDECRTIPEACRGDMCV 60
Db 1 MPGIKRLITVITIALCLPSPGNAOACTNGFDLDRSGCCLDIDECRTIPEACRGDMCV 60
QY 61 NONGGYLCIPRTNPNVYRGPSNPSTYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
Db 61 NONGGYLCIPRTNPNVYRGPSNPSTYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
QY 121 ESNQCVDBECATDSHOCNPTQICINTEGGYTCSTGDTGWLLEGGCLDIDECRYGCAQL 180
Db 121 ESNQCVDBECATDSHOCNPTQICINTEGGYTCSTGDTGWLLEGGCLDIDECRYGCAQL 180
QY 181 CANVPGSYSTCTCNGPFTLNEDGRSCQDVNECATENPCVQTCVNTYGSFICRCDPGELEE 240
Db 181 CANVPGSYSTCTCNGPFTLNEDGRSCQDVNECATENPCVQTCVNTYGSFICRCDPGELEE 240
QY 241 DGVHCSMDMECSSEFLCOHECVNQPGTYFCSCPFGYILLDDNRSCQDINECEHNRHNCN 300
Db 241 DGVHCSMDMECSSEFLCOHECVNQPGTYFCSCPFGYILLDDNRSCQDINECEHNRHNCN 300
QY 301 LOOTCYNLOGGFRCIDPICEEPLRISDNRCMCPAENPGCRDOPFTILYRDMDVSGRS 360
Db 301 LOOTCYNLOGGFRCIDPICEEPLRISDNRCMCPAENPGCRDOPFTILYRDMDVSGRS 360

PRIOR APPLICATION NUMBER: 09/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380138
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380139
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/403296
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/403297
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/423741
PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: 09/423844
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 09/522342
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: 09/548815
PRIOR FILING DATE: 2000-04-13
PRIOR APPLICATION NUMBER: 09/664610
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: 09/665350
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: 09/709238
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 09/767609
PRIOR FILING DATE: 2001-01-22
PRIOR APPLICATION NUMBER: 09/802706
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 09/808689
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 09/866028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 09/870574
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: 09/872035
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 09/886342
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: PCT/US98/14552
PRIOR FILING DATE: 1998-07-14
PRIOR APPLICATION NUMBER: PCT/US98/18824
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: PCT/US98/19093
PRIOR FILING DATE: 1998-09-14
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: PCT/US98/19437
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: PCT/US98/24855
PRIOR FILING DATE: 1998-11-20
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: 1998-12-01
PRIOR APPLICATION NUMBER: PCT/US98/25190
PRIOR FILING DATE: 1998-11-25
PRIOR APPLICATION NUMBER: PCT/US99/05028
PRIOR FILING DATE: 1999-03-08
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/20111
PRIOR FILING DATE: 1999-09-01
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547

Query Match 99.8%; Score 2527; DB 9; Length 448;

Best Local Similarity 99.8%; Pred. No. 2,1e-164; Mismatches 1; Indels 0; Gaps 0;

Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MFGIRILVTTLALCLPSPGNAAOCTNGFDLDRSGGCLDIDECRTIPACRGDMCV 60
DB 1 MFGIRILVTTLALCLPSPGNAAOCTNGFDLDRSGGCLDIDECRTIPACRGDMCV 60

QY 61 NONGGILCIPRTNPVRYGPNPSTPYSGPPAABPLSAPNPTISRPLICRGYQMD 120
DB 61 NONGGILCIPRTNPVRYGPNPSTPYSGPPAABPLSAPNPTISRPLICRGYQMD 120
QY 121 ESNOCVDVDECATDSHOCNPOTICINTEGGVTCSCDTGMYMLEGQCCLDIDECRGGYCOOL 180
DB 121 ESNOCVDVDECATDSHOCNPOTICINTEGGVTCSCDTGMYMLEGQCCLDIDECRGGYCOOL 180
QY 181 CANVPSSYSTCNPFTLNEDGNSCODVNECATENPCVQTCVNTYGSFICRDPGELEE 240
DB 181 CANVPSSYSTCNPFTLNEDGNSCODVNECATENPCVQTCVNTYGSFICRDPGELEE 240
QY 241 DGVHCSMDSCSSEFLCOHECVNQPCTYSCSPGTTLLDNNRSCODINCEHRNHTCN 300
DB 241 DGVHCSMDSCSSEFLCOHECVNQPCTYSCSPGTTLLDNNRSCODINCEHRNHTCN 300
QY 301 LOOTCYNLOGGFCIPDIRCEEPYLRISDRNCPAENPCGRDPTILYRDMVSGRS 360
DB 301 LOOTCYNLOGGFCIPDIRCEEPYLRISDRNCPAENPCGRDPTILYRDMVSGRS 360
QY 361 VPADIFOMQATTPYGAATYIFQIKSGNEGEFTWROTGPISATLWTRPIKPREIQLDL 420
DB 361 VPADIFOMQATTPYGAATYIFQIKSGNEGEFTWROTGPISATLWTRPIKPREIQLDL 420
QY 421 EMITVTVINFRGSSVIRLRITVSOYPF 448
DB 421 EMITVTVINFRGSSVIRLRITVSOYPF 448

RESULT 6
US-10-174-590-408

Sequence 408, Application US/10174590
Publication No. US20030008352A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Pan, James

APPLICANT: Smith, Victoria

APPLICANT: Watanabe, Colin K.

APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE OF INVENTION: ACIDS ENCODING THE SAME

FILE REFERENCE: P3430R1C42

CURRENT APPLICATION NUMBER: US/10/174,590

CURRENT FILING DATE: 2002-06-18

Prior application removed - See file wrapper or Palm

NUMBER OF SEQ ID NOS: 612

SEQ ID NO 408

LENGTH: 448

TYPE: PRT

ORGANISM: Homo Sapien

US-10-174-590-408

Query Match 99.8%; Score 2527; DB 9; Length 448;

Best Local Similarity 99.8%; Pred. No. 2,1e-164; Mismatches 1; Indels 0; Gaps 0;

Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MFGIRILVTTLALCLPSPGNAAOCTNGFDLDRSGGCLDIDECRTIPACRGDMCV 60
DB 1 MFGIRILVTTLALCLPSPGNAAOCTNGFDLDRSGGCLDIDECRTIPACRGDMCV 60
QY 61 NONGGILCIPRTNPVRYGPNPSTPYSGPPAABPLSAPNPTISRPLICRGYQMD 120
DB 61 NONGGILCIPRTNPVRYGPNPSTPYSGPPAABPLSAPNPTISRPLICRGYQMD 120
QY 121 ESNOCVDVDECATDSHOCNPOTICINTEGGVTCSCDTGMYMLEGQCCLDIDECRGGYCOOL 180
DB 121 ESNOCVDVDECATDSHOCNPOTICINTEGGVTCSCDTGMYMLEGQCCLDIDECRGGYCOOL 180

Oy 361 VPADIFQMATTTPYRGAYIIFQIKSGNEGREFYMTGPISATLVMTPTIKSPREIOLDL 420
Db 361 VPADIFQMATTTPYRGAYIIFQIKSGNEGREFYMTGPISATLVMTPTIKSPREIOLDL 420
Oy 421 EMITVTNFINFRGSSVIRLRIYVSQYRF 448
Db 421 EMITVTNFINFRGSSVIRLRIYVSQYRF 448

RESULT 5
US-10-066-500-15
; Sequence 15, Application US/10066500
; Patent No. US20020177165A1
; GENERAL INFORMATION:
; APPLICANT: Avi J. Ashkenazi
; APPLICANT: Kevin P. Baker
; APPLICANT: David A. Bolstein
; APPLICANT: Luc Desnoyers
; APPLICANT: Dan L. Eaton
; APPLICANT: Napoleone Ferrara
; APPLICANT: Sherman Fong
; APPLICANT: Wei-Qiang Gao
; APPLICANT: Hanspeter Gerber
; APPLICANT: Mary E. Gerltsen
; APPLICANT: Audrey Goddard
; APPLICANT: Paul J. Godowski
; APPLICANT: Austin L. Gurney
; APPLICANT: Ivar J. Klyavin
; APPLICANT: Jennie P. Mather
; APPLICANT: Mary A. Napier
; APPLICANT: James Pan
; APPLICANT: Nicholas F. Paoni
; APPLICANT: Margaret Ann Roy
; APPLICANT: Timothy A. Stewart
; APPLICANT: Daniel Tumas
; APPLICANT: Colin K. Watanabe
; APPLICANT: P. Mickey Williams
; APPLICANT: William I. Wood
; APPLICANT: Zemin Zang
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P130R1C7
; CURRENT APPLICATION NUMBER: US/10/066, 500
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 10/002,796
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062285
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063733
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066840
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 60/069694
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 60/074086
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074092

; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/095998
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/097000
; PRIOR FILING DATE: 1998-08-18
; PRIOR APPLICATION NUMBER: 60/099601
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099803
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099811
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099812
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/100858
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/101922
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/106032
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/109304
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/125778
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 60/139695
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 60/145070
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/149396
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/169495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 08/918874
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 08/933821
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 08/960507
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 09/114844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: 09/136801
; PRIOR FILING DATE: 1998-08-19
; PRIOR APPLICATION NUMBER: 09/136804
; PRIOR FILING DATE: 1998-08-19
; PRIOR APPLICATION NUMBER: 09/136828
; PRIOR FILING DATE: 1998-08-19
; PRIOR APPLICATION NUMBER: 09/158342
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: 09/180997
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 09/202088
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: 09/254311
; PRIOR FILING DATE: 1999-03-03
; PRIOR APPLICATION NUMBER: 09/254460
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: 09/254465
; PRIOR FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: 09/284663
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 09/332928
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: 09/332929
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: 09/333075
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: 09/333077
; PRIOR FILING DATE: 1999-06-14

STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/275,805
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/839,525
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: PF224
REFERENCE/DOCKET NUMBER: 36,373
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8509
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 448 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-275-805-2

Query Match 100.0%; Score 2533; DB 10; Length 448;
Best Local Similarity 100.0%; Pred. No. 8.3e-165;
Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 MPGIRLITVTLALCLPSGNAQAQCTNGFDLRSGQCLDIDECRTIPEACRGDMCV 60
QY 1 NONGGYLCIPRTNPYRGPSNPYSTPYSGPYPAAPPLSAPNYPPTISRLICRFGYOMD 120
DB 61 NONGGYLCIPRTNPYRGPSNPYSTPYSGPYPAAPPLSAPNYPPTISRLICRFGYOMD 120
QY 61 NONGGYLCIPRTNPYRGPSNPYSTPYSGPYPAAPPLSAPNYPPTISRLICRFGYOMD 120
DB 61 NONGGYLCIPRTNPYRGPSNPYSTPYSGPYPAAPPLSAPNYPPTISRLICRFGYOMD 120
QY 121 ESNOCVADVDECATDSHOCNPOTICINTEGGYTCTDGYWLLGGQCLDIDECRGYCOOL 180
DB 121 ESNOCVADVDECATDSHOCNPOTICINTEGGYTCTDGYWLLGGQCLDIDECRGYCOOL 180
QY 181 CANPGSYSCNPGFTLNEDGRSCODVNECATENPCVOTCVMTYGSFICRCPGYELE 240
DB 181 CANPGSYSCNPGFTLNEDGRSCODVNECATENPCVOTCVMTYGSFICRCPGYELE 240
QY 241 DGVHCSMDDECSFSEFLQHECVNPGTYFCSCPFGYILLDDNRSCODINECHRNHTCN 300
DB 241 DGVHCSMDDECSFSEFLQHECVNPGTYFCSCPFGYILLDDNRSCODINECHRNHTCN 300
QY 301 LOQTCTYNOGGFKCIDPRICEEPYLRISDNRMCAPENPGCRDOPFTILYRDMDVYSGRS 360
DB 301 LOQTCTYNOGGFKCIDPRICEEPYLRISDNRMCAPENPGCRDOPFTILYRDMDVYSGRS 360
QY 361 VPAIFQMAQATTRYPGAYYIFQIKSGNEGREFYMRGTGPISATLVMTRP1KGREIQLDL 420
DB 361 VPAIFQMAQATTRYPGAYYIFQIKSGNEGREFYMRGTGPISATLVMTRP1KGREIQLDL 420
QY 421 EMITVNTVINFRGSSVIRLRIYVSQYPF 448
DB 421 EMITVNTVINFRGSSVIRLRIYVSQYPF 448

RESULT 4
US-09-836-561-1
Sequence 1, Application US/09836561
Patent No. US20020038006A1

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
Corley, Neil C.
Guegler, Karl J.
TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/836,561
FILING DATE: 16-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/212,168
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0333 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 448 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: CORNNOT01
CLONE: 45517
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-836-561-1
Query Match 100.0%; Score 2533; DB 10; Length 448;
Best Local Similarity 100.0%; Pred. No. 8.3e-165;
Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 MPGIRLITVTLALCLPSGNAQAQCTNGFDLRSGQCLDIDECRTIPEACRGDMCV 60
QY 1 NONGGYLCIPRTNPYRGPSNPYSTPYSGPYPAAPPLSAPNYPPTISRLICRFGYOMD 120
DB 61 NONGGYLCIPRTNPYRGPSNPYSTPYSGPYPAAPPLSAPNYPPTISRLICRFGYOMD 120
QY 61 NONGGYLCIPRTNPYRGPSNPYSTPYSGPYPAAPPLSAPNYPPTISRLICRFGYOMD 120
DB 61 NONGGYLCIPRTNPYRGPSNPYSTPYSGPYPAAPPLSAPNYPPTISRLICRFGYOMD 120
QY 121 ESNOCVADVDECATDSHOCNPOTICINTEGGYTCTDGYWLLGGQCLDIDECRGYCOOL 180
DB 121 ESNOCVADVDECATDSHOCNPOTICINTEGGYTCTDGYWLLGGQCLDIDECRGYCOOL 180
QY 181 CANPGSYSCNPGFTLNEDGRSCODVNECATENPCVOTCVMTYGSFICRCPGYELE 240
DB 181 CANPGSYSCNPGFTLNEDGRSCODVNECATENPCVOTCVMTYGSFICRCPGYELE 240
QY 241 DGVHCSMDDECSFSEFLQHECVNPGTYFCSCPFGYILLDDNRSCODINECHRNHTCN 300
DB 241 DGVHCSMDDECSFSEFLQHECVNPGTYFCSCPFGYILLDDNRSCODINECHRNHTCN 300
QY 301 LOQTCTYNOGGFKCIDPRICEEPYLRISDNRMCAPENPGCRDOPFTILYRDMDVYSGRS 360
DB 301 LOQTCTYNOGGFKCIDPRICEEPYLRISDNRMCAPENPGCRDOPFTILYRDMDVYSGRS 360

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 448 amino acids
TYPE: amino acid
STRANDEDNESS: No. US20020165151A1 Relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-041-016-2

Query Match 100.0%; Score 2533; DB 9; Length 448;
Best Local Similarity 100.0%; Pred. No. 8.3e-165;
Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGKIRLLVTITIALCLPSGNAOACTNGDLDROSQCCLDIDECRTIPKCRDMMCV 60
DB 1 MGKIRLLVTITIALCLPSGNAOACTNGDLDROSQCCLDIDECRTIPKCRDMMCV 60
QY 61 NONGYLICIPRTNPNVYRGPSNPNSTPSGYPAPAAPLSAPNPTISRPLICRGYQMD 120
DB 61 NONGYLICIPRTNPNVYRGPSNPNSTPSGYPAPAAPLSAPNPTISRPLICRGYQMD 120
QY 121 ESNQCVADVDECAIDSHQCNPTQICINTEGGYTCSDTGWMLLEGQCLDIDECRYGCOOL 180
DB 121 ESNQCVADVDECAIDSHQCNPTQICINTEGGYTCSDTGWMLLEGQCLDIDECRYGCOOL 180
QY 181 CANVPGSYSTCTNPGFTLNEDGRSCODVNECATENPCVQTCVNTYGSFTICRDPGELEE 240
DB 181 CANVPGSYSTCTNPGFTLNEDGRSCODVNECATENPCVQTCVNTYGSFTICRDPGELEE 240
QY 241 DGVHCSMDCECSFSEFLCOHECVNQPCTYFSCPPGYILLDDNRSCODINECEHNNHCN 300
DB 241 DGVHCSMDCECSFSEFLCOHECVNQPCTYFSCPPGYILLDDNRSCODINECEHNNHCN 300
QY 301 LQOTCYNLOGGFKCIDIRCEEPYLRISDNRCMCPAENPGCRDQFTLLYRDMDVSGRS 360
DB 301 LQOTCYNLOGGFKCIDIRCEEPYLRISDNRCMCPAENPGCRDQFTLLYRDMDVSGRS 360
QY 361 VPADIFOMATTRTPGAYYIFQIKSGNEGREFYMRQTPISATLVMTPIKPREIQLDL 420
DB 361 VPADIFOMATTRTPGAYYIFQIKSGNEGREFYMRQTPISATLVMTPIKPREIQLDL 420
QY 421 EMITVNTVINFRGSSVIRLRIVYSQYPF 448
DB 421 EMITVNTVINFRGSSVIRLRIVYSQYPF 448

RESULT 2
US-09-083-002-2
Sequence 2, Application US/09083002
Patent No. US2001001650A1

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Racine, Lisa A.
APPLICANT: LaValle, Edward R.
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Evans, Cheryl
APPLICANT: Agostino, Michael
APPLICANT: Lu, Zhijian
APPLICANT: Honjo, Tasuku
APPLICANT: Tashiro, Kei
APPLICANT: Nakamura, Tomoyuki
TITLE OF INVENTION: SECRETED PROTEINS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083.002
FILING DATE:
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: P-41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 448 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-083-002-2

Query Match 100.0%; Score 2533; DB 10; Length 448;
Best Local Similarity 100.0%; Pred. No. 8.3e-165;
Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGKIRLLVTITIALCLPSGNAOACTNGDLDROSQCCLDIDECRTIPKCRDMMCV 60
DB 1 MGKIRLLVTITIALCLPSGNAOACTNGDLDROSQCCLDIDECRTIPKCRDMMCV 60
QY 61 NONGYLICIPRTNPNVYRGPSNPNSTPSGYPAPAAPLSAPNPTISRPLICRGYQMD 120
DB 61 NONGYLICIPRTNPNVYRGPSNPNSTPSGYPAPAAPLSAPNPTISRPLICRGYQMD 120
QY 121 ESNQCVADVDECAIDSHQCNPTQICINTEGGYTCSDTGWMLLEGQCLDIDECRYGCOOL 180
DB 121 ESNQCVADVDECAIDSHQCNPTQICINTEGGYTCSDTGWMLLEGQCLDIDECRYGCOOL 180
QY 181 CANVPGSYSTCTNPGFTLNEDGRSCODVNECATENPCVQTCVNTYGSFTICRDPGELEE 240
DB 181 CANVPGSYSTCTNPGFTLNEDGRSCODVNECATENPCVQTCVNTYGSFTICRDPGELEE 240
QY 241 DGVHCSMDCECSFSEFLCOHECVNQPCTYFSCPPGYILLDDNRSCODINECEHNNHCN 300
DB 241 DGVHCSMDCECSFSEFLCOHECVNQPCTYFSCPPGYILLDDNRSCODINECEHNNHCN 300
QY 301 LQOTCYNLOGGFKCIDIRCEEPYLRISDNRCMCPAENPGCRDQFTLLYRDMDVSGRS 360
DB 301 LQOTCYNLOGGFKCIDIRCEEPYLRISDNRCMCPAENPGCRDQFTLLYRDMDVSGRS 360
QY 361 VPADIFOMATTRTPGAYYIFQIKSGNEGREFYMRQTPISATLVMTPIKPREIQLDL 420
DB 361 VPADIFOMATTRTPGAYYIFQIKSGNEGREFYMRQTPISATLVMTPIKPREIQLDL 420
QY 421 EMITVNTVINFRGSSVIRLRIVYSQYPF 448
DB 421 EMITVNTVINFRGSSVIRLRIVYSQYPF 448

RESULT 3
US-09-275-805-2
Sequence 2, Application US/09275805
Patent No. US20010051358A1

GENERAL INFORMATION:

APPLICANT: OLSEN, HENRIK S.
APPLICANT: LI, HAODONG
TITLE OF INVENTION: EXTRACELLULAR EPIDERMAL GROWTH FACTOR
TITLE OF INVENTION: LIKE PROTEIN
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 17:53:42 ; Search time 88.4684 Seconds
(without alignments)
582.319 Million cell updates/sec

Title: US-09-674-379a-13

Perfect score: 2533

Sequence: 1 MGKIRILVTILALCPSP.....INRGSSVIRLRIVSQYPF 448

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 440863 seqs, 114992915 residues

Total number of hits satisfying chosen parameters: 440863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published_Applications-AA:*

1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2533	100.0	448	9	US-10-041-016-2
2	2533	100.0	448	10	US-09-083-002-2
3	2533	100.0	448	10	US-09-275-805-2
4	2533	100.0	448	10	US-09-836-561-1
5	2527	99.8	448	9	US-10-066-500-15
6	2527	99.8	448	9	US-10-174-590-408
7	2527	99.8	448	9	US-10-176-758-408
8	2527	99.8	448	9	US-10-175-737-408
9	2527	99.8	448	9	US-10-173-706-408
10	2527	99.8	448	9	US-10-173-738-408
11	2527	99.8	448	9	US-10-175-752-408
12	2527	99.8	448	9	US-10-176-482-408
13	2527	99.8	448	9	US-10-176-757-408
14	2527	99.8	448	9	US-10-176-913-408
15	2527	99.8	448	9	US-10-180-552-408
16	2527	99.8	448	9	US-10-180-557-408
17	2527	99.8	448	9	US-10-173-700-408
18	2527	99.8	448	9	US-10-174-572-408
19	2527	99.8	448	9	US-10-174-579-408

20	2527	99.8	448	9	US-10-174-582-408	Sequence 408, App
21	2527	99.8	448	9	US-10-174-588-408	Sequence 408, App
22	2527	99.8	448	9	US-10-175-739-408	Sequence 408, App
23	2527	99.8	448	9	US-10-175-740-408	Sequence 408, App
24	2527	99.8	448	9	US-10-175-743-408	Sequence 408, App
25	2527	99.8	448	9	US-10-176-488-408	Sequence 408, App
26	2527	99.8	448	9	US-10-176-492-408	Sequence 408, App
27	2527	99.8	448	9	US-10-176-747-408	Sequence 408, App
28	2527	99.8	448	9	US-10-176-750-408	Sequence 408, App
29	2527	99.8	448	9	US-10-176-985-408	Sequence 408, App
30	2527	99.8	448	9	US-10-176-987-408	Sequence 408, App
31	2527	99.8	448	9	US-10-176-991-408	Sequence 408, App
32	2527	99.8	448	9	US-10-176-993-408	Sequence 408, App
33	2527	99.8	448	9	US-10-184-658-408	Sequence 408, App
34	2527	99.8	448	9	US-10-002-796-15	Sequence 15, App1
35	2527	99.8	448	9	US-10-066-494-15	Sequence 15, App1
36	2527	99.8	448	9	US-10-173-695-408	Sequence 408, App
37	2527	99.8	448	9	US-10-173-697-408	Sequence 408, App
38	2527	99.8	448	9	US-10-173-705-408	Sequence 408, App
39	2527	99.8	448	9	US-10-174-576-408	Sequence 408, App
40	2527	99.8	448	9	US-10-174-585-408	Sequence 408, App
41	2527	99.8	448	9	US-10-174-585-408	Sequence 408, App
42	2527	99.8	448	9	US-10-174-586-408	Sequence 408, App
43	2527	99.8	448	9	US-10-174-586-408	Sequence 408, App
44	2527	99.8	448	9	US-10-176-481-408	Sequence 408, App
45	2527	99.8	448	9	US-10-176-481-408	Sequence 408, App

ALIGNMENTS

RESULT 1
US-10-041-016-2
; Sequence 2, Application US/10041016
; Patent No. US20020165151A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
McCoy, John M.
Racie, Lisa A.
Lavallee, Edward R.
Merberg, David
Treacy, Maurice
Evans, Cheryl
Agostino, Michael
Lu, Zhijian
Honjo, Tasuku
TITLE OF INVENTION: SECRETED PROTEINS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSER: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/041,016
FILING DATE: 07-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/083,002
FILING DATE: 21-MAR-1998
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: P-41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851



1
2

Db 663 MTORGCEIDDECLNPS-TC-PDEQCVNPSGYCPCBTBGRMGNGCLDYDECLPNV 720
QY 177 CQO-LEANPGASISCTCNPGFTLNEDEGRSCQDVNECAPVQ----- 219
Db 721 CANGCDNLEGSYMCCHKGYTRTPDHKCRDIDECQOGLNVCNGCKNTGSEFRCTGQ 780
QY 220 -----TCVNTGSEFCRCDBPGYLEEDGVHCSMDCECF 253
Db 781 GYOLSAKQCEIDECQHRHLCAHGOCNTEGSEFCVCDQGYRASSGLDHCEDINECLE 840
QY 254 SEFLCOH-BCVNPQTYFSCPCPGYILLDDNRSCQDINECEHRNHTCNLOQTCYNLOGCF 312
Db 841 DKSVCGRCQDINAGSYDCTCPDGF-QLDDNKTCQDINECEHPG-LGGPQCECLNTGSEF 898
QY 313 KCI-----DPRCEPYLRIS-----DN-----RCMC-----PAENPGCR 342
Db 899 HCVCQGGFISADGRTCEIDECVNTVCDSHGFCNTAGSFRLCYGQAPQDGGCV 958
Y 343 DQFTILYRDMDVSG 358
Db 959 D-----VNECELLSG 968

RESULT 12
US-08-897-443-1
Sequence 1, Application US/08897443
Patent No. 5981263

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

APPLICANT: Lal, Preeti

APPLICANT: Corley, Neil C.

APPLICANT: Shah, Purvi

TITLE OF INVENTION: HUMAN MATRILIN-3

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/897,443

FILING DATE: Filed Herewith

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PE-0348 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 638 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: UTRSN0702

CLONE: 681719

US-08-897-443-1

Query Match

Best Local Similarity

18.88; Score 475; DB 2; Length 638;

28.48; Pred. No. 5.4e-30;

Matches 96; Conservative 61; Mismatches 117; Indels 64; Gaps 11;
QY 15 ICLPSPGNAOCTNTEGFDLDRSGOCLDIECTIPACGDMACVONGICLIPRNP 74
Db 293 LCYVNPVGSVCCYCYSTALAEDEGRCAVDYCASENHGEHE--CVNADSYLC----- 344
QY 75 YVRGPYSNPYSPPYSGPYAPAPLAPNPTLSRPLICRFGYOMD-ESNOCVVDDECAT 133
Db 345 -----QCHGEGALNNDKTKTCYIDVC 367
QY 134 DSHOCNPVICINTEGTYGCTDGYWLEGO--CLDIDECRYG--CQOLCANVPGSYS 189
Db 368 SNHGQ--HECVNTDYSYCHCLKPTLPDCKTKCRINFCALNKPCEHECVNMESSY 425
QY 190 CTCNPGFTLNEDEGRSCQDVNECA-TENPCVQTCVNTGYSFICRCDGTELEEDGVHCSMD 248
Db 426 CCHRGYITLDPNGKTSRVDHCAQDHDGECQLNTEDESFVCCSEGLTNEDEKTESRV 485
QY 249 DECSFSEFLQHECVNQPETYFSCPCPGYILLDDNRSCQDINECEHRNHTCNLOQTCYNL 308
Db 486 DYCLSDHGCCEYSVNMDSFACQCEGHVLRSDGKCAKALDSCALGDHGC--EHSVSS 543
QY 309 OGFEKCIDPRCEPYLRISDNR-C-----MCPAENPGC 341
Db 544 EDSFVC-----QCFEGITILREDGKTCRRKQVCOAIDHGC 577

RESULT 13

US-08-479-722B-2

Sequence 2, Application US/08479722B

Patent No. 6074840

GENERAL INFORMATION:

APPLICANT: Bonadio, Jeffrey

APPLICANT: Yin, Mushan

TITLE OF INVENTION: LATENT TGF β BINDING PROTEIN (LTBP)

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Williams, Morgan & Amerson

STREET: 7676 Hillmont, Suite 250

CITY: Houston

STATE: Texas

COUNTRY: USA

ZIP: 77040

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/479,722B

FILING DATE: 07-JUN-1995

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US PCT/US95/02251

FILING DATE: 21-FEB-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/316,650

FILING DATE: 30-SEP-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/199,780

FILING DATE: 18-FEB-1994

ATTORNEY/AGENT INFORMATION:

NAME: Fussey, Shelley P.M.

REGISTRATION NUMBER: 39,458

REFERENCE/DOCKET NUMBER: 4100, 000500/EUS

TELECOMMUNICATION INFORMATION:

TELEPHONE: (713) 934-7011

TELEFAX: (713) 934-7011

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 183 amino acids

TYPE: amino acid

```

NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/212,168
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/884,072
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0333 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ. ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 458228
US-09-212-168-5

Query Match
Best Local Similarity 48.3%; Score 963.5; DB 4; Length 387;
Matches 171; Conservative 55; Mismatches 123; Indels 5; Gaps 4;

QY 96 APPLSAPNPTTISRPLICRGYOMDESNOVCVDECATDSHQCNPTQICINTEGGYTCSC 155
DB 38 ABOQRIPNRP--SHRIQCAAGYEGSEHNVCODIDECTAGTINCGRADYQICINLRGSFACQC 95
QY 156 TGGYMLLEGGCLIDERTY-CYQQQLCANYPGYSCTCNPFITLINEDGRSCODVNECATE 214
DB 96 PRGYQRGEGCVVIDECTIPYCHQRGVNTPGSEFYCCQSPGFOLANNVTCVDINECDAS 155
QY 215 NPCVQTCVNTYGSFICRDPGELEEDGVHCSMDSECSFSEFLCOHECVNOPGTGYEFCSC 274
DB 156 NCAQCCYNILGTFICOCNGOYELSSDRINCEIDDECTSSYLCOYQCVNEFGKSCMCP 215
QY 275 PGYILLDNRSQDINECEHRNMTCNLQQTCTYNLOGGFKCIDPICEEPYLRISDNRCMC 334
DB 216 OGYYVY-RSRFTCODINECEFTTNE-CRDEMCWMYHGFRCYPRNRCOPYLITPENRCVC 273
QY 335 PAENPCGROPTILVRDMVVSGRSVPADIFQWATTRYGAYTIFQIKSNEGREFRM 394
DB 274 PVSNAACRELPQSIYKYSIRSDRSVSDIFQIQAATTIVANTINTFRIKSNEGGEFYL 333
QY 395 ROTGPISATLVMPRIKPREIQLDLEMTVNTVINFGSSVIRLRIYVSQYPF 448
DB 334 RQTSFVSAMLYVYKLSGPREHIVDLEMTVYSIGTFRITSSVLRLLITVGPSPF 387

RESULT 10
US-09-188-930-336
; Sequence 336, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:

```

```

APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
APPLICANT: Murison, James Greg
TITLE OF INVENTION: Compositions Isolated From Skin Cells
FILE REFERENCE: 11000.1011c1
CURRENT APPLICATION NUMBER: US/09/188,930A
CURRENT FILING DATE: 1998-11-09
NUMBER OF SEQ. ID NOS: 348
SOFTWARE: FASTSEQ for windows Version 3.0
SEQ. ID NO 336
LENGTH: 274
TYPE: PR1
ORGANISM: Human
US-09-188-930-336

Query Match
Best Local Similarity 36.0%; Score 912.5; DB 4; Length 274;
Matches 154; Conservative 47; Mismatches 73; Indels 1; Gaps 1;

QY 174 YGYCQQLCANYPGYSCTCNPFITLINEDGRSCODVNECATENPCVQTCVNTYGSFICRD 233
DB 1 YRYCQHRGVNLPQSFRCQCEPFGOLGPNRSCVDNECDMGAPCEQRCFNSYGTFLRCRH 60
QY 234 PGYELEEDGVHCSMDSECSFSEFLCOHECVNOPGTGYEFCSCPGYITLLDNRSQDINECE 293
DB 61 QGYELHNDGFSQSIDECSSYLCQYRCVNEPFGFSCHCPQGYOLL-ATRLCQDIDEEB 119
QY 294 HRNHTCNLQQTCTYNLOGGFKCIDPICEEPYLRISDNRCMCPAENPCGROPTILYRDM 353
DB 120 SGAHQCEAQTCVNFHGGYRCVDNRCVEPYIQVSENRCLCPASNPLCREQSSIVHRYM 179
QY 354 DVYSGRSVPADIFQWATTRYGAYTIFQIKSNEGREFMYOTGPISATLVMPRIKRP 413
DB 180 TITERSVPAVDVFOIATSVYPGAYNAFOIRNGSQGDFYIROIINVSAMLVLRVYGP 239
QY 414 REIQLDEMTVNTVINFGSSVIRLRIYVSQYPF 448
DB 240 REYVLDEMTVNTVMSGRASSVLRITVFGAYTF 274

RESULT 11
5177197-30
Patent No. 5177197
APPLICANT: KANZAKI, TETSURO; OLOFSSON, ANDERS; MOREN, ANITA;
WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAESSON-WELSH,
LENA; HELDIN, CARL-HENRIK
TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING
HUMAN TRANSFORMING GROWTH FACTOR-BETA1-BINDING PROTEIN
NUMBER OF SEQUENCES: 53
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/487,343
FILING DATE: 27-FEB-1990
SEQ. ID NO: 30
LENGTH: 1394
5177197-30

Query Match
Best Local Similarity 21.5%; Score 544; DB 6; Length 1394;
Matches 128; Conservative 52; Mismatches 134; Indels 122; Gaps 17;

QY 16 CLPSPGNAQOCATGFDLDRSGGLDIDERTYPEACRGDMVCVNONGYLCTPRTNV 75
DB 562 CINLPVRYTICIGYRFSEQRKQCVVIDECTQVQHLG-SQRCENTBESSFLCI----- 614
QY 76 YRGPEYSNPYSTPYSGPYPAAPPLSAPNPTISRPLI-----CRGTYQ 118
DB 615 -----CPAGFMASEEGTNCIDVDECLRPDYCGEGHCYNTVGAFAFCEYCDSGYR 662
QY 119 MDESNCQVDVDECATDTHQCNPTQICITBEGYTC-SCDGYWMLLEGGCLDIDEC-RYCY 176

```

ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEO for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980,514
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0436 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEO ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:

US-08-980-514-3

Query Match	38.0%;	Score 963.5;	DB 3;	Length 387;
Best Local Similarity	48.3%;	Pred. No. 5.8e-69;		
Matches 171; Conservative	55;	Mismatches 123;	Indels 5;	Gaps 4

QY 96 APPISAPNYPITSRPLICRFGYQMNDESNOCVDVDECATDSHOCPITOCINTEGGYTCS 155

156 TDGWWLLEGGGLDIDECRRY-gyccoolcANvpgsYscTcnpGFTLlNEdGRScODvNECATE 214
D5 ADPQGLPSNP---SHRIQCAAGYqgSESHNVCQDIDECTAGTNCRADVCINLqGSFACC 95

Db 96 PPGIOKRGEOCDIDECTIPPYCHORCVNTPGSGFYCQCSPGFLAANNYYTCVDINECDAS 155

[illegible]

275 PGYLLDNRSCODINECEHNRNHCNLAOTCYNLOGGFKCIDPIRCEEPYLRISDNRMC 334

Db 216 QGYGVV-RSPTQDINECETINE-CREDEMCWNHGGFRCPRNPCDDYILPENRCVC 273

274 PVSNAACRELPQSIYVKYMSIRSDRSVPDIFQIQAATTIYANTINFRIRKSGENGEFPL 333

[illegible]

US-09-212-168-5
: Sequence 5, Application US/09212168

* Filing No.: 0507100
 * GENERAL INFORMATION:
 * APPLICANT: Bandman, Olga

APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS

```

APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0436 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-845-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 443 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRSTN013
CLONE: 2786449
US-08-980-514-1

```

Query Match 50.9%; Score 1289; DB 3; Length 443;

Best Local Similarity 49.9%; Pred. No. 1,1e-94; Matches 226; Conservative 74; Mismatches 131; Indels 22; Gaps 5;

```

QY 1 MPKIRILVTYIALCLPSPGNAQ-----AOCINGFDLRSGGCLDIDECRTPEACRG 55
DB 8 LPSLLMALLLLLGASAPDSSEPDSTECTDGYEMDPSQHCRDYNELTPEACKG 67
QY 56 DMACVNGNGYLCIPRTNPPYRGYSNPSTIPYAAAPPLSAPNPTISRPLICRF 115
DB 68 EMKCMHNGGYLCIPRSAAVINDLHG-----EGP-PPYPFPAOHPPN-----PCPP 111
QY 116 GYCMDSNOCVDVDECATDSHOCNPTOICINTEGGYTSCDITGWLLEGGCLDIDECR 175
DB 112 GYEPDDQSDVDVDECAALHDCRPSQDCNHLPSGYCTCPDGKRIKIGECVDIDECR 171
QY 176 YCOQLCANVPGSSYCTCNPFTLNEDGRSCDVNCCATENPCVOTCVNTYGSFICRDPG 235
DB 172 YCHRCVNLPSFRCQCEPFGOLGPNPNNSCVDVNCMDGAPCEGCFNSYGFELCRCHG 231
QY 236 YLEBEGVHSDMDECFSEFLCOHECVNOGRTYFCSPPGTILLDDNRSCODINECR 295
DB 232 YELHRCGFCSDIDECSSYSLCYRCVNEBGRFSCHPGQGLL-ARRLCODIDECESG 290
QY 296 NHTCNLOQTCYNLOGGFKCIDPICEEPYLRISDRNCMPAENPCGROPTILYRDMV 355
DB 291 AHCCSEAGTCVNRHGRICVDYRNCVEPIYGVSEKCLCPASNPICREPPSIVHRYMTI 350
QY 356 VSGRVPADIFOMATRRYPGAYYIFQIKSGNEGHEFYMRGTPISATLVMTRPKIGPRE 415
DB 351 TSSRVPADYFOIATSVYFGAYNAFOIRAGNSQODFYIROINNVSAMLVLARPTGPRE 410
QY 416 IQDLEMTYNTYINFRGSSYIRLRIVSOYCF 448
DB 411 YVLDLEMTNLSMYSRASSVLRITVFAGATF 443

```

RESULT 6
US-08-884-072-5
Sequence 5, Application US/08884072

```

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA

```

```

ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/884,072
FILING DATE: Herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0333 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-845-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 458228
US-08-884-072-5

```

Query Match 38.0%; Score 963.5; DB 2; Length 387;

Best Local Similarity 48.3%; Pred. No. 5.8e-69; Matches 171; Conservative 55; Mismatches 123; Indels 5; Gaps 4;

```

QY 96 APPLSAPNPTISRPLICRFYQMDNESQCVYDVECATDSHOCNPTOICINTEGGYTSCG 155
DB 38 ADPRIRSNP--SHRIGCAAGYEGSEHNVQDIDECTAGTHNCRADQVCINLRGSPACG 95
QY 156 TDGWLLEGGCLDIDECRY-GYCOQLCANVPGSSYCTCNPFTLNEDGRSCDVNCCAT 214
DB 96 PPGYQKRGECQVDIDECYIPYCHQRCVNTP6SFYCCSPGFLANNATCYDINCCDS 155
QY 215 NPCVQTCVNTYGSFICRCDPGELEEDGVHCSMDMECSFSEFLCOHECVNOGRTYCSGP 274
DB 156 NQCAQCYNTLIGSTICQNOGYELSSDRLNCEIDECRTSYLCQTCVNEPEKSCMCP 215
QY 275 PGYILLDNNSCODINECEHNRHTCNLOQTCYNLOGGFKCIDPICEEPYLRISDRNCM 334
DB 216 QGYGVV-RRSRQDINEETNE-CREDEMCMNHGFRCYRNPQODPYILPEKRCVC 273
QY 335 PAENPGCRDQFTLLYRDMVYSGRVPADIFOMATRRYPGAYYIFQIKSGNEGHEFYM 394
DB 274 PVSNAMEKRELPOSIVYKXMSIRSDRSPSDIFQIATTIYANTINTFRKSENENGEFYL 333
QY 395 ROTGPISATLVMTRPKIGPREIOLDLEMTYNTYINFRGSSYIRLRIVSOYCF 448
DB 334 ROTSPVAMLVVKSLSGPREHYVLDLEMTYSSIGTFRTSSVLRITIVGPF 387

```

RESULT 7
US-08-833-963C-9

Sequence 9, Application US/08833963C

```

GENERAL INFORMATION:
APPLICANT: Olsen, et al.
TITLE OF INVENTION: Extracellular/Epidermal Growth Factor
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESSES:
ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Ave

```

RESULT 5
 US-08-980-514-1
 Sequence 1, Application US/08960514
 Patent No. 6004753
 GENERAL INFORMATION:
 APPLICANT: Yue, Henry
 APPLICANT: Guegler, Karl J.
 APPLICANT: Shah, Purvi
 TITLE OF INVENTION: HUMAN SI-5-ECMP-LIKE PROT
 TITLE OF INVENTION: EIN
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Dr.
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/980,514
 FILING DATE: Filed Herewith
 PRIOR APPLICATION DATA:

Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPGIKRLVTITIALCLPSGNAOACTNGFDLROSQCCLIDECRTIPEACRDMCY 60
DB 1 MPGIKRLVTITIALCLPSGNAOACTNGFDLROSQCCLIDECRTIPEACRDMCY 60
QY 61 NONGYLICIPRTNPNVYRGPSNPSYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
DB 61 NONGYLICIPRTNPNVYRGPSNPSYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
QY 121 ESNQCVYDECAIDSHOQCNPTQICINTEGGYTCSTGTYWLLGQCCLIDECRYGCOOL 180
DB 121 ESNQCVYDECAIDSHOQCNPTQICINTEGGYTCSTGTYWLLGQCCLIDECRYGCOOL 180
QY 181 CANVPSYSTCNGPFTLNEDGRSCODVNECATENPCVQTCVNTYGSFICRCDPYELEE 240
DB 181 CANVPSYSTCNGPFTLNEDGRSCODVNECATENPCVQTCVNTYGSFICRCDPYELEE 240
QY 241 DGVHCSMDSECSFSEFLCOHECVNQPGTYFCSPPGYILLDDNRSQCCLIDECRHNHCN 300
DB 241 DGVHCSMDSECSFSEFLCOHECVNQPGTYFCSPPGYILLDDNRSQCCLIDECRHNHCN 300
QY 301 LOOTCYNLOGGFICIDIRCEEPYLRISDNRCMCPAENPGCRDOPFTILYRDMVYSGRS 360
DB 301 LOOTCYNLOGGFICIDIRCEEPYLRISDNRCMCPAENPGCRDOPFTILYRDMVYSGRS 360
QY 361 VPADIFQMATTTRPGAYYIFQIKSGNEGREFYMRQTPISATLVMTRPKIPREIQLDL 420
DB 361 VPADIFQMATTTRPGAYYIFQIKSGNEGREFYMRQTPISATLVMTRPKIPREIQLDL 420
QY 421 EMITVNTVINFRGSSVIRLRIVYSQYPF 448
DB 421 EMITVNTVINFRGSSVIRLRIVYSQYPF 448

RESULT 2
US-09-212-168-1
Sequence 1, Application US/09212168
Patent No. 6303765
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/212,168
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/84,072
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0333 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 448 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: CORNNOT01
CLONE: 45517
US-09-212-168-1

Query Match 100.0%; Score 2533; DB 4; Length 448;
Best Local Similarity 100.0%; Pred. No. 2,6e-193;
Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPGIKRLVTITIALCLPSGNAOACTNGFDLROSQCCLIDECRTIPEACRDMCY 60
DB 1 MPGIKRLVTITIALCLPSGNAOACTNGFDLROSQCCLIDECRTIPEACRDMCY 60
QY 61 NONGYLICIPRTNPNVYRGPSNPSYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
DB 61 NONGYLICIPRTNPNVYRGPSNPSYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
QY 121 ESNQCVYDECAIDSHOQCNPTQICINTEGGYTCSTGTYWLLGQCCLIDECRYGCOOL 180
DB 121 ESNQCVYDECAIDSHOQCNPTQICINTEGGYTCSTGTYWLLGQCCLIDECRYGCOOL 180
QY 181 CANVPSYSTCNGPFTLNEDGRSCODVNECATENPCVQTCVNTYGSFICRCDPYELEE 240
DB 181 CANVPSYSTCNGPFTLNEDGRSCODVNECATENPCVQTCVNTYGSFICRCDPYELEE 240
QY 241 DGVHCSMDSECSFSEFLCOHECVNQPGTYFCSPPGYILLDDNRSQCCLIDECRHNHCN 300
DB 241 DGVHCSMDSECSFSEFLCOHECVNQPGTYFCSPPGYILLDDNRSQCCLIDECRHNHCN 300
QY 301 LOOTCYNLOGGFICIDIRCEEPYLRISDNRCMCPAENPGCRDOPFTILYRDMVYSGRS 360
DB 301 LOOTCYNLOGGFICIDIRCEEPYLRISDNRCMCPAENPGCRDOPFTILYRDMVYSGRS 360
QY 361 VPADIFQMATTTRPGAYYIFQIKSGNEGREFYMRQTPISATLVMTRPKIPREIQLDL 420
DB 361 VPADIFQMATTTRPGAYYIFQIKSGNEGREFYMRQTPISATLVMTRPKIPREIQLDL 420
QY 421 EMITVNTVINFRGSSVIRLRIVYSQYPF 448
DB 421 EMITVNTVINFRGSSVIRLRIVYSQYPF 448

RESULT 3
US-09-188-930-186
Sequence 186, Application US/09188930A
Patent No. 6150502
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
TITLE OF INVENTION: Compositions Isolated From Skin Cells
FILE REFERENCE: 11000.1011C1
CURRENT APPLICATION NUMBER: US/09/188,930A
NUMBER OF SEQ ID NOS: 348
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 186
LENGTH: 337
TYPE: PRT
ORGANISM: Rat
FEATURE:
NAME/KEY: UNSURE
LOCATION: (14)...(124)
NAME/KEY: UNSURE
LOCATION: (135)...(135)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 17:54:52 ; Search time 13.8875 Seconds
(without alignments)
949.162 Million cell updates/sec

Title: US-09-674-379a-13
Perfect score: 2533
Sequence: 1 MFGIRILVTILALCLPSP.....INFRGSSVIRLRIVSYQYF 448

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfillseq1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2533	100.0	448	2	US-08-884-072-1
2	2533	100.0	448	4	US-09-212-168-1
3	1791	70.7	337	4	US-09-188-930-186
4	1289	50.9	443	2	US-08-833-963C-2
5	1289	50.9	443	3	US-08-833-963C-2
6	963.5	38.0	387	2	US-08-884-072-5
7	963.5	38.0	387	2	US-08-833-963C-9
8	963.5	38.0	387	3	US-08-980-514-3
9	963.5	38.0	387	4	US-09-212-168-5
10	912.5	36.0	274	4	US-09-188-930-336
11	544	21.5	1394	6	517197-30
12	475	18.8	638	2	US-08-897-443-1
13	463	18.3	1833	3	US-08-479-722B-2
14	463	18.3	1833	5	PCT-US95-02251-18
15	461	18.2	1251	5	PCT-US95-02251-3
16	461	18.2	1252	1	US-08-199-780-3
17	461	18.2	1252	1	US-08-316-650-3
18	461	18.2	1253	2	US-08-479-722B-4
19	452	17.8	956	2	US-08-897-443-3
20	392	15.5	886	3	US-09-110-116-3
21	377	14.9	2471	1	US-08-185-432-16
22	377	14.9	2471	1	US-08-083-590A-19
23	377	14.9	2471	3	US-08-532-384-19
24	377	14.9	2471	4	US-08-899-232-1
25	375	14.8	810	2	US-08-820-170A-34
26	375	14.8	810	3	US-09-055-699-34
27	375	14.8	810	4	US-09-273-565-34

28	375	14.8	810	4	US-09-565-538-34	Sequence 34, Appl
29	375	14.8	810	4	US-09-661-468-34	Sequence 34, Appl
30	366	14.4	816	2	US-08-820-170A-37	Sequence 37, Appl
31	366	14.4	816	3	US-09-055-699-37	Sequence 37, Appl
32	366	14.4	816	4	US-09-273-565-37	Sequence 37, Appl
33	366	14.4	816	4	US-09-565-538-37	Sequence 37, Appl
34	366	14.4	816	4	US-09-661-468-37	Sequence 37, Appl
35	360.5	14.2	1964	4	US-09-467-997-1	Sequence 1, Appl
36	359	14.2	2703	1	US-08-185-432-19	Sequence 19, Appl
37	359	14.2	2703	4	US-08-899-232-4	Sequence 4, Appl
38	350.5	13.8	652	2	US-08-751-305-2	Sequence 2, Appl
39	347.5	13.7	2556	1	US-08-083-590A-20	Sequence 20, Appl
40	347.5	13.7	2556	3	US-08-532-384-20	Sequence 20, Appl
41	346	13.7	2523	1	US-08-185-432-18	Sequence 18, Appl
42	346	13.7	2523	4	US-08-899-232-3	Sequence 3, Appl
43	344.5	13.6	2556	1	US-08-185-432-17	Sequence 17, Appl
44	344.5	13.6	2556	4	US-08-899-232-2	Sequence 2, Appl
45	340.5	13.4	673	1	US-08-282-141-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-884-072-1
; Sequence 1, Application US/08884072
; Patent No. 5872234
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/884,072
; FILING DATE: Herewith
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0333 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 448 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: CORNMO101
; CLONE: 45517
; US-08-884-072-1
Query Match 100.0%; Score 2533; DB 2; Length 448;
Best Local Similarity 100.0%; Pred. No. 2.6e-193;

Db 785 P1ECNPGYTLGSDKCVLIDECOKONGCS--HRCSTNGSGFKC----SCPPGYELDSQ 838
QY 331 R-CM---CPAENPGC 341
Db 839 KTCODIDECDDDKTSC 854

RESULT 15

096SC3 PRELIMINARY; PRT: 2673 AA.
AC 096SC3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Fibulin-6 (Fragment).
GN FIBL-6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
NC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MELANOMA;
RA Kostka G., Timpl R.;
RT "Partial sequence of fibulin-6 with a C-terminal region related to
domain II and III of the fibulin family."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ306906; CAC37630.1; -;
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000875; Cectropin.
DR InterPro: IPR000561; EGF-Like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00008; EGF_5.
DR Pfam: PF00047; Ig_17.
DR Pfam: PF00090; TSP_1; 6.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_5.
DR PROSITE: PS00268; CECROPIN; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; UNKNOWN_3.
DR PROSITE: PS01187; EGF_CA; UNKNOWN_8.
DR PROSITE: PS50092; TSP1; 6.
FT NON_TER 1
SQ SEQUENCE 2673 AA; 291017 MW; BEAEC30B8340E272 CRC64;

Query Match 21.6% Score 548; DB 4; Length 2673;
Best Local Similarity 26.9%; Pred. No. 7.4e-46;

Matches 144; Conservative 66; Mismatches 203; Indels 122; Gaps 17;

QY 16 CLPSPGNAQACTNGFPLDRSQGLDIDECRTIPEACRGDMGVNONGYLCTIPRTNPV 75
Db 2159 CHNMGTYTSCPKGLTIADGRICQDIDECALGRHCHAGODCDNTIGSTRVRCGSG 2218
QY 76 YRGPTSNPSTPYSGPYPAAPPLSADNPYPTISRPILICRFGYQMDENQCVADVDECATDS 135
Db 2219 FRRT-SDGLSCODINEQESSPCHQRCFNAIGSFHCCEPGYQL-KGRKCMADVNECR--Q 2274
QY 136 HQCNPOICINTEGGYTC--SCTDGYWLE-GCGLDIDECRYCQ----QLCANVPGSY 188
Db 2275 NVCRPDCHKNTGSGYCIDLCPNGMTKAENGTCIDIDECRGTQHCRCRYNQCENRTGSY 2334
QY 189 SCTCNGFTLNEGRSCDVNEC-ATENPCVQTCVNTYGSFICRDPGYLEEDVHCS- 246
Db 2335 KVCYPRGYSRQYGRPCMDINECRQVPRCAHQCSNTPGSFKCICPPGHLIGDCKSCAG 2394
QY 247 ----- 246
Db 2395 LERLPVNGTQYSSYNLARFSVRNNYQPOQHRYQYSHLYSSYSEYRNSRTSLSTRRTIR 2454
QY 247 -----DMDECSFSEFLCOHECVNQPGTYFCSPPEGYILLDDNRSCQDINE 291
Db 2455 KTCPEGSEASHDRCVDDIDECENTD-ACQHECKNTFGSYOCICPPGYQLTHNGKTCQDIDE 2513

QY 292 CEHRNHTCMLOQTCYNLGGFKCIDPICEPYLRIS-DNRKM--CPAENPGCDOPFTI 348
Db 2514 CLEQNVHCGPNRMCFFNRRSGYQCID-TPCPPNYORDVSGFCLKNCPNDLECALSPYAL 2572
QY 349 LYRDMDVSGRYPADIFQMOATR---YGAAYI-----FOIKSGNREGRPYMR 395
Db 2573 EYKLVSLPFGIATNODLIRLVAAYTODGVHMRPTFTFLAVDEQTVPFALRDN----- 2624
QY 396 QTPISATLVMTRPIKGR--EIQDLEMITVNTVINFRGSSVRLRIYVSOYFE 448
Db 2625 ----LKGVVYTTIRPLLEAEYTRMRVRASSYSANGTIERQYTFIV--YIANSAPY 2673

Search completed: July 3, 2003, 18:21:57
Job time : 35.9041 secs

```

OY 69 --IP--RTNPYRGP---YSNPYRPGPYPAAPP-----98
DB 227 RLVRHRRANRIGNARRRRRMDPYR--AGEYRASQANTFEGCPMGHLPQHGHCVDIDE 284
OY 99 -----LSAPNPTISRPLICRFQYOMD---ESNOCV-----126
DB 285 CATLMDDCLESORCLNTPGSGFKCIRLSCGTYAMDESETERNNCFLLIINTFNCKYFEV 344
OY 127 -DVDECATSHOCNPTQICINTEGTYTCS-----154
DB 345 EVDVDECNLASHDCGLPYQCRNTGSGYRKDAKCKGDELQNMTEGYIDEVYTHNGAGE 404
OY 155 -----CTDGYWL--LEGOCLDIDECRYGCOOL-CANPGYSCTGNP 194
DB 405 ECVNTPGSRCCQKGMLCAGYEYVNGATGFCEDVNECQGVCSMECINLFGTYCKRCKGP 464
OY 195 GFTLNE-----DGRSCQDVNE 210
DB 465 GYEFNDAKRCEVDDECIRKAGHYCDLSACINTIGSFECKCRPGFOLASDGRCEDEVNE 524
OY 211 CATE-NPCVQTCVNTGSGFICRCDPGELEEDGVHCSDMBCSF-----SEFLQHECVNQ 265
DB 523 CTTGIAACBQKCVNIGSYQICIDRGFALPDGTRKCEIDECSTWAGSGNDLGMGCGINT 584
OY 266 PGTYSFCSPPGYILLDNRSQDINECEHNNHCNLAQTCYNLGGFKCIDPICEEPLYL 325
DB 585 KCSYLQCCPFGYKIQPDGRVCYVDECA-NGECAGSKVCYNTIGSFRC-HSIDCPYTI 642
OY 326 RISDN-----CMCPAENFCGRDQPTIYRDMDVYSGRSVP-----362
DB 643 HSLNNRCNRPSACGLPEE---CSKVPLFTFYQISL-ARAVPISHRPAITLFKYS 697
OY 363 ----ADI---FQMGATTRYGATYIYQIKSGNEGREYMTGP--ISATLVMPPIGP 413
DB 638 APNMDTEYVNFELQKTITVYAGNVLPALIRAN-----FLQKGKRNNAVVTLRSLDGP 752
OY 414 REIOLDL 420
DB 753 QTVKIQDL 759

```

RESULT 14

Q9VS89 PRELIMINARY; PRT; 1394 AA.

AC Q9VS89: 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DT CG7526 protein (Fragment).

OS Drosophila melanogaster (Fruit fly)

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydriidae; Drosophilidae; Drosophila.

OC NCBI_Taxid=7227;

RA BRUNSON R.C., ROGERS Y.H.C., BLAZER R.G., CHAMPE M., PFEIFFER B.D.,

RA WANG K.H., DOYLE C., BAXTER E.G., HELT G., NELSON C.R., MIKLOS G.L.G.,

RA ABRIEL J.F., ABBAYANI A., AN H.-Y., ANDREWS-PINAKOCH C., BALDWIN D.,

RA BALLEW R.M., BASU A., BAKENDALE J., BAYRAKTARGLU L., BEASLEY E.M.,

RA BEESON K.Y., BENOS P.V., BERMAN B.P., BHANDARI D., BOLSHAKOV S.,

RA BOROVKA D., BOTCHAN M.R., BOUCK J., BROKSTEIN P., BROTLIER P.,

RA BURTLIS K.C., BUSAM D.A., BUTLER H., CADIEU E., CENTER A., CHANDRA I.,

RA CHERRY J.M., CAWLEY S., DAHLKE C., DAVENPORT L.B., DAVIES P.,

RA DE PABLOS B., DELCHER A., DENG Z., MAYS A.D., DEW I., DIETZ S.M.,

```

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.J., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RA "The genome sequence of Drosophila melanogaster."
RA Science 287:2185-2195(2000).
RA EMBL: AE003558; AAF50538.1; -.
RA HSSP: P00736; IAP0.
RA Flybase: FBgn0035798; CG7526.
RA InterPro: IPR000152; ASX_Hydroxyl.
RA InterPro: IPR000561; EGF_Like.
RA InterPro: IPR001881; EGF_CA.
RA InterPro: IPR003410; Hyalin.
RA InterPro: IPR000436; Sush1_SCR_CCP.
RA InterPro: IPR001491; Thrombomodulin.
RA Pfam: PF00008; EGF_11.
RA Pfam: PF00084; sush1_2.
RA PRINTS: PR00907; THROMBOMODULN.
RA SMART: SM00132; CCP_2.
RA SMART: SM00179; EGF_CA_9.
RA SMART: SM00001; EGF_Like_5.
RA PROSITE: PS00010; ASX_HYDROXYL_8.
RA PROSITE: PS01186; EGF_2_9.
RA PROSITE: PS01187; EGF_CA_10.
RA Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
RA NON_TER 1394 1394
RA SEQUENCE 1394 AA; 152269 MW; CD29380E3162F68A CRC64;

```

Query Match 22.0%; Score 558.5; DB 5; Length 1394;

Best local similarity 32.4%; Pred. No. 3e-47;

Matches 122; Conservative 42; Mismatches 135; Indels 77; Gaps 13;

```

OY 15 LCIPSPGNAOCTNGRD---LDROSGCLDIDECRTIPENACRGMMCVNONGYLICPR 71
DB 507 VCNRLPSSYSGICAGAGELKLDIRGCFDIDECSTQTHGSDMLLENLSTYCL-- 564
OY 72 TNPNYRGPYNPSTPYSGPYPAAPLAPNPTISRPLICRFQYOMDESNOQVAVDVC 131
DB 565 CPFGYALGLDNHVTSLNSSFITDSTSETTS-----AATCIDIDEC 606
OY 132 ATDSDHOCNPTQICINTEGTYTCSCTDGYWLEGO--CIDIDEC--RYGYCOOLCANVPS 187
DB 607 SLANGNC--SHFQCNREGGFGQACPLGVALSEDMRTCODIDECIDLSNQCQSLCINQPG 664
OY 188 YSTCTNPGFTLNDGRSCQVNECATE-NPCVQTCVNTGSGFICRCDPGELEEDGVHCS 246
DB 665 FACACETGFEFTDGFECACDIDECSDYGNCSDCITINLGHACGCEGYELAKKLSCL 724
OY 247 DMEDEC-SFSEFLQHECVNPGTYFCSPPGYILLDNRSQD-----287
DB 725 DVDECAGLSLSGSGSHEDINNAGTFCGCPILGIIINDGRSSCPALVGPGRQASADGA 784
OY 288 -----DINECEHNNHCNLAQTCYNLGGFKCIDPICEEPLYRISDN 330

```

```

DB 387 GNDLCMGCGINTGSGYLCCGCPGPKIOPDGTCTVDVDECA MGECAGSDKCVNTLGSFK 445
OY 314 CIPDICEEPLYLISDRN-----C-MCPAENPC-RDQPTTILYRDMDVYSGRSV- 361
DB 446 C-HSIDCPTNYIHDLSLNKNOIADGYSICIKVCTEDTCLGNHTREVIYQFAVPSLKII 504
OY 362 -PADI-----FOMQATTRYPGAYVYFQIKSGNEGREFYKRGCPISATILVTRPI 410
DB 505 SPIEVSRIYTHMGVPPSVYDNLDIYGGHRHRIVOERNIG-----IVQLVKPI 551
OY 411 KGPREFOLDLEMTVTNTVINFSGSVIR-----LRIVSOYPF 448
DB 552 SGP-----TVEIKVINHTKSRFTGVILAFNEAIIISVSKYF 589

RESULT 12
OY 090H16 PRELIMINARY; PRT; 554 AA.
DB 01-MAY-2000 (TREMBlrel. 13, Created)
DB 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DB 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DB DJ162H14.1 (Fibulin 1) (Fragment).
GN FBLN1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lloyd D.;
RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z68047; CAB62995.1; -.
DR HSSP; P35555; IEMN.
DR InterPro; IPR000020; Anaphylatoxin.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR Pfam; PF01821; ANATO; 3.
DR Pfam; PF00008; EGF; 6.
DR SMART; SM00104; ANATO; 3.
DR SMART; SM00179; EGF_Ca; 4.
DR SMART; SM00001; EGF-like; 5.
DR PROSITE; PS01177; ANAPHYLATOXIN_1; 3.
DR PROSITE; PS01178; ANAPHYLATOXIN_2; 3.
DR PROSITE; PS00010; ASX_HYDROXYL; 4.
DR PROSITE; PS01186; EGF_2; 3.
DR EGF-like domain; Glycoprotein; Hydroxylation.
FT NON TER
SQ SEQUENCE 554 AA; 59767 MW; FE285184599A2982 CRC64;

Query Match 23.0%; Score 583; DB 4; Length 554;
Best Local Similarity 36.1%; Pred. No. 3.4e-50;
Matches 116; Conservative 47; Mismatches 102; Indels 56; Gaps 13;

OY 16 CLPSPGNAOACTNGFDLDRSGOCLDIDECRTTPEACRGDMCVNONGYLCTPRTNPV 75
DB 261 CRP-----KLCCKGSGFLID-ALGNCIGINECLISAPCPHGTCTINTGSSYTC----- 307
OY 76 YRGPYNSPYSPYSGPYPAAPRISAPYPTISRLTRFQGM-DESNQCVVDDECATD 134
DB 308 -----QKNVPR-----CGRGYHNEEGTRCVGVDECAPR 336
OY 135 SHOCNPTQICINTEGTYCTSDGYWL-LEGOCLDIDEC-RY-GYCOOLCANVPGSVS 189
DB 337 AEPCCGKRCVNSPGRCECKTGYFDGISMVGVNCECRYPORLCHKCENTLGLSTL 396
OY 190 CTGNPPTLLEDGRSCODVNECATENPCVQTCVNTYSGFICRCDGYELE-DEYHCSDM 248
DB 397 CSCSVGERLSDVRSCGINECSS-SPSCDECANVYSGCYCRGQYGLSDVDGYTCBGI 455
OY 249 DECSF--SEPLCQHECVNQPRTYFCSCP-PGYILLDNRSCODINECHRNHTCNLQOTC 305

```

```

DB 456 DECALETPTGHCISYRCINIPGSCSPSSGYRLAPGRNCGIDECVTGIHNSINETC 515
OY 306 YNLGGFCRCDIPRCEPYLR 326
DB 516 FNIQGGFRCL-APECPENYR 535

RESULT 13
OY 018026 PRELIMINARY; PRT; 798 AA.
DB 018026
AC 018026: Q20903;
DB 01-JAN-1998 (TREMBlrel. 05, Created)
DB 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DB 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DB F56H11.1 protein.
GN F56H11.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodertinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilkinson J.;
RL Submitted (JAN-1996) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkhen R.,
RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaughan M., Vaughan K., Waterson R.,
RA Watson A., Weinstein L., Wilkinson-Sproat J., Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
RN [3]
RP SEQUENCE FROM N.A.
RA Lloyd C.;
RL Submitted (DEC-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z68749; CAA92962.1; -.
DR EMBL; Z68219; CAA92962.1; JOINED.
DR EMBL; Z68219; CAA92483.1; -.
DR EMBL; Z68749; CAA92483.1; JOINED.
DR HSSP; P16109; IFSB.
DR InterPro; IPR000020; Anaphylatoxin.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR Pfam; PF01821; ANATO; 2.
DR Pfam; PF00008; EGF; 5.
DR SMART; SM00104; ANATO; 2.
DR SMART; SM00179; EGF_Ca; 4.
DR SMART; SM00001; EGF-like; 6.
DR PROSITE; PS01177; ANAPHYLATOXIN_1; UNKNOWN_1.
DR PROSITE; PS00010; ASX_HYDROXYL; 4.
DR PROSITE; PS01186; EGF_2; 5.
DR PROSITE; PS01187; EGF_Ca; 8.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 798 AA; 87205 MW; 3BF1EE9BD54D8BF9 CRC64;

Query Match 22.8%; Score 577.5; DB 5; Length 798;
Best Local Similarity 26.7%; Pred. No. 1.9e-49;
Matches 162; Conservative 50; Mismatches 180; Indels 215; Gaps 23;

OY 15 LCLPSPG-NAOACTNGFDLDRSGOCLDIDECRTTPEACRGDMCVNONGYLCTPRTNPV 68
DB 167 LCHDRGERVEYSCRSRGFDLAPDGMACVDRNECLITRQSPCTQSDCVNTIGTYICQRRIS 226

```

Db 472 VFERMGSSAVPGD-----SMOAGHRRQ 495

RESULT 10

ID 042182 PRELIMINARY; PRT; 681 AA.

AC 042182;

DT 01-JAN-1998 (TREMBLrel. 05, Created)

DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Fibulin-1 D.

GN FBLN1.

OS Brachydanio rerio (zebrafish) (Zebra danio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;

OC Cyprinidae; Danio.

OX NCBI_TaxID=7955;

RN (1)

RP SEQUENCE FROM N.A.

NA Zhang H.-Y., Lardelli M., Ekblom P.;

NA "Sequence of zebrafish fibulin-1 and its expression in developing

NA heart and other embryonic organs."

RL Dev. Genes Evol. 0:0-0(1997).

DR EMBL; AF013751; AAB80944.1; -.

DR HSSP; P35555; 1EMN.

DR ZFIN; ZDB-GENE-990415-73; fbln1.

DR InterPro; IPR000020; Anaphylatoxin.

DR InterPro; IPR000152; Asx_hydroxyl.

DR InterPro; IPR000561; EGF-like.

DR Pfam; PF01821; ANATO; 1.

DR SMART; SM00008; EGF; 4.

DR SMART; SM00104; ANATO; 2.

DR SMART; SM00179; EGF_CA; 5.

DR SMART; SM00001; EGF_Like; 4.

DR PROSITE; PS01177; ANAPHYLATOXIN_1; UNKNOWN_1.

DR PROSITE; PS01178; ANAPHYLATOXIN_2; 1.

DR PROSITE; PS00010; ASX_HYDROXYL; 3.

DR PROSITE; PS01186; EGF_2; 3.

DR PROSITE; PS01187; EGF_CA; 6.

KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.

SEQUENCE 681 AA; 74459 MW; 175C966305A46699 CRC64;

Query Match 24.9%; Score 631; DB 13; Length 681;

Best Local Similarity 32.6%; Pred. No. 6.3e-55;

Matches 153; Conservative 65; Mismatches 162; Indels 90; Gaps 22;

16 CLPSPGNAOCTNGFDLDRSGOCLDIDECRTIPENACRGDMGVNONGYLICIPRTNPV 75

265 CRP-----RMQCAAGFTID-ALGSCIDINECVSTALSRG-QMCFNTVGSFTICQHS--- 314

76 YRGPTSNPTSTYSGYPPAAPPLSAPNPTISRPLICRGYQMD-ESNOCVDVDECATD 134

315 -----VTCGRGYHLNAGTRCVDIDECAGP 339

135 SHOCNPICINTEGGYTCSTGYWL--LEGCLDIDECRGY-----COOLANVPSPY 188

340 DNSCD-GHCICINLVSTRCERTGFTFNSTSRSCEDIDECR-NYGRILCAHKCEIILASY 397

189 SCTCNPGFTLNDGRSCODVNECATENPCVQTCVNTGSEFCRCDPGYELEEE-DGVHCS 247

398 KCSCTAGFLADGRNCDVNECES-SPCSQGCANVGSYOSYCRGQSLDADGITEED 456

248 MDECF--SEFLCAHCVNPGTYFCSCP-PGYIILDNRSCODINECEHNNHTCNLAQT 304

457 IDECALPTGSHICSYCHNTPGSFHCTCPASGYTLANGRSQCIDIDCLTGHSCSESES 516

305 CYNLAGGFCICPIRCEEPYLRISDN-----RCM--CPAENGCRRDQPTIIVRD 352

517 CFNIGGFCILS-FDCPRANTRRSQDTRPRVADRADIIRCAVSCQNDISCVLP--ILSHS 573

353 MDVVS-----GRSVPADIFOMQATT-----RYPGAYYIFQIKSGNEGREFYMQT--GP 399

Db 574 HTAISLPTRFREFKPREIYVLRSPTRPHLPHMDSPEIYDILEGNIONSFIIRKLDHGM 633

QY 400 ISATLVMPRIKPPREIQLDLEM-ITYVTVINFGSSYIRKRIYVSQTF 448

Db 634 IGVVAKVRLVGPVRLTKLAMYVNGVSHR--NINVIYVSEWF 681

RESULT 11

ID 097251 PRELIMINARY; PRT; 589 AA.

AC 097251;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Fibulin-1D (Fragment).

GN FBLN1.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN-CB1489 HIM-8;

RC MEDLINE-99120531; Pubmed-9923656;

RA Barth J.L., Argaves K.M., Roark E.F., Little C.D., Argaves W.S.;

RT "Identification of the C. elegans fibulin-1 gene."

RT characterization of the C. elegans fibulin-1 gene."

Matrix Biol. 17:635-646(1998).

DR EMBL; AF070477; AAC24035.1; -.

DR HSSP; P16109; IFSB.

DR InterPro; IPR000152; Asx_hydroxyl.

DR InterPro; IPR000561; EGF-like.

DR InterPro; IPR001881; EGF_CA.

DR Pfam; PF00008; EGF; 6.

DR SMART; SM00179; EGF_CA; 5.

DR SMART; SM00001; EGF_Like; 4.

DR PROSITE; PS00010; ASX_HYDROXYL; 4.

DR PROSITE; PS01186; EGF_2; 5.

DR PROSITE; PS01187; EGF_CA; 7.

KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.

FT NON_TER 1

SEQUENCE 589 AA; 63984 MW; 8EA3E8FCE0B97BE6 CRC64;

Query Match 23.5%; Score 596; DB 5; Length 589;

Best Local Similarity 27.3%; Pred. No. 1.8e-51;

Matches 159; Conservative 53; Mismatches 163; Indels 208; Gaps 21;

15 ICLPSPG-NAOACTNGFDLDRSGOCLD-IDECRTIPENACRGDMGVNONGYLICIPRT 72

66 LCHDRGGEKEKSCRSRSGFDLAPDGMACVDHIDECATLMDCLDSQRCLNTGSEFKCI--- 122

73 NPVYRGPTSNPTSTYSGYPPAAPPLSAPNPTISRPLICRGYQMD-ESNOCVDVDECATD 131

123 -----RTLSCGTGYAMDSERECRDVDEC 146

132 ATDSDHCNPICINTEGGYTC-----SCTDGYLLEGGCLDID 170

147 NLSHDCGPLYCRRNTOGSTRCDAKKCGDGLQNPMTGECTSYCPNIGYTPKNGKCNDID 206

171 ECRVGY-----COO--LCA-- 182

207 ECTTGNGCAGGECVNTPGSFRCQKKNLCAHGEVNGATGFCEDVNECQGVCSMBCI 266

183 NPGYSTCNCNPGFTLNE----- 200

267 NLPGTAKCKGPGYEFNDAKKCEDVDECIFAGHVDLSAECINTIGSFCKCKPGRQL 326

201 --DGRSCODVNECATENPCVQTCVNTGSEFCRCDPGYELEEDGVHCSMDECSF----- 253

327 ASDGRCEDEVNETTGTGACCEQKCVNTPGSTQCLCDKRFALGPBGTKEDEIDECSTWAGS 386

254 SEFLCAHCVNPGTYFCSCP-PGYIILDNRSCODINECEHNNHTCNLAQTQCYNLQGGFK 313

Db 539 FNETCNIQGGFRLS-LECPENRYKSGDVFLEKTDITRCIKSRPNVNCVLDPVHTI 597
 QY 349 LYRDMVSGRSV--PADIFOMQ--TRRYG--AYIFQISGNEGREF-----YMRGTG 398
 Db 598 SRIVISLPTRETRRETFRELLFLRAITPTYPANQADIIFFDITEGRLRESFDIIRKM--DG 655
 QY 399 PISATLWMPRIPIKGPRIQDLEM-ITVNTVYINFRGSSVIRLRIYVSQYPF 448
 Db 656 MTGVGVROYRPIGVFPFHALIKLENNYMGVSHR--NIYVNHLEFSEYWF 704

RESULT 8

Q9Y3V7 PRELIMINARY; PRT; 576 AA.
 ID Q9Y3V7;
 AC Q9Y3V7;
 DT 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE Hypothetical 63.3 kDa protein (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=UTERUS;
 RA Wambutt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;
 RL Submitted (Mar-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL050095; CAB43267.1; -
 DR HSSP: P00736; IAPQ.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF_Ca.
 DR Pfam: PF00008; EGF_6.
 DR SMART: SM00179; EGF_CA; 8.
 DR SMART: SM00001; EGF_Like; 2.
 DR PROSITE: PS00010; ASX_HYDROXYL; 4.
 DR PROSITE: PS01186; EGF_2; 4.
 DR PROSITE: PS01187; EGF_CA; 9.
 DR Calcium-binding: EGF-like domain; Glycoprotein; Hydroxylation;
 KW Hypothetical protein; Repeat.
 FT NON_TER 1
 SO SEQUENCE 576 AA; 63274 MW; FF2F0E9B185D8AC3 CRC64;

Query Match 28.1%; Score 711.5; DB 4; Length 576;
 Best Local Similarity 34.3%; Pred. No. 3.9e-63;
 Matches 146; Conservative 61; Mismatches 156; Indels 63; Gaps 12;

QY 15 LCLSPGN---AQACGTNGFDLDRSGGCLDIDECRTPEACRGDMGVNONGYLCTP 70
 Db 172 LCONTKGSFYCOARORCMDF-LDDEGNCVDINECTSLSEPCRPFGSCINTVGSYTC-- 228
 QY 71 RTNPNYRGPNSPYSTPSGPPAAPPLISAPNYPISRLICRFYQ--MDESNQCVND 129
 Db 229 -----QNRPLICARVHASDDTKCVND 252
 QY 130 ECATDSHQCNPTQICINTGGYCTSDGYW--LLEGCLDIDECRYG---VCOQLCANV 184
 Db 253 ECEGVHRSRGVYCHNLPGSYRCCKAGFORDAFGRCIDVNECHASGRCLQHCNCE 312
 QY 185 PGASISCHNPGFTLNEDGRSCODVNECATENPCVQTCVNTYFSFICRCDPGYELEBDYH 244
 Db 313 LGSYRSCASGFLLAADKRCEDVNECAQR--CSQECANITYSGYCYCRQYQALADGHR 371
 QY 245 CSDMDECS--FSEFLQHECVNQPGTYFCSCP--PGYILLDDNRSQDINCEHRNHTCNQ 302
 Db 372 CTQIDECAGAGGILCTFRCLVNPSSYQACPEQGTTPRANGSRCKDVCALGTHNCSBA 431
 QY 303 QTCYNLGGFKCIDPIRCEEPYLRISDNRCMCAPENPCRD-----QPTLLYRMDVY 356
 Db 432 ETCNIGGSFRL--RFECPPNYVGVSKTKC-----ERTCHDLPLECONSARLTHQNLQ 486

QY 357 SGRSVADPIFOMQATTRYPGAYYIFQIKSGNEGREFYMRGTGPISATLWMPRIKGPRI 416
 Db 487 TGLVPAHIFRIGPAPAFGTDTALNTIKNGNEBGFSTRRLNATGVVYIQRAVLEPRDF 546
 QY 417 OLDLEM 422
 Db 547 ALDVEH 552

RESULT 9

Q9HBO5 PRELIMINARY; PRT; 495 AA.
 ID Q9HBO5
 AC Q9HBO5;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Hypothetical 54.3 kDa protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gu J.R., Man D.F., Zhao X.T., Zhou X.M., Jiang H.Q., Zhang P.P.,
 RA Qiu W.X., Huang Y., Qiu X.K., Qian L.F., He L.P., Li H.N., Yu Y.,
 RA Yu J., Han L.H.;
 RT "Novel Human cDNA clones with function of inhibiting cancer cell
 growth."
 RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF217999; AAG17241.1; -
 DR HSSP: P35555; IEMN.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF_Ca.
 DR Pfam: PF00008; EGF_6.
 DR SMART: SM00181; EGF_9.
 DR SMART: SM00179; EGF_CA; 9.
 DR SMART: SM00001; EGF_Like; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; 3.
 DR PROSITE: PS01186; EGF_2; 3.
 DR PROSITE: PS01187; EGF_CA; 8.
 KW EGF-like domain; Glycoprotein; Hydroxylation; Hypothetical protein.
 SO SEQUENCE 495 AA; 54340 MW; C40434E6C82ED3D70 CRC64;

Query Match 26.7%; Score 676; DB 4; Length 495;
 Best Local Similarity 35.8%; Pred. No. 1.3e-59;
 Matches 139; Conservative 59; Mismatches 128; Indels 62; Gaps 16;

QY 16 CLPSPGNAQACGTNGFDLDRSGGCLDIDECRTPEACRGDMGVNONGYLCTPNTYV 75
 Db 158 CRP-----KLOCKSGFTQD-ALGNCLIDINECLISAPCPIGHGTCINTBSYTC----- 204
 QY 76 YRGPSNPNYSTPSGPPAAPPLISAPNYPISRLICRFYQ--MDESNQCVND 134
 Db 205 -----QKNVNP-----CGRGVHLEBGRVNDVDECA 233
 QY 135 SHQCNPTQICINTGGYCTSDGYW--LLEGCLDIDECRYG---VCOQLCANV 189
 Db 234 AEPGCKHRCVNSPSSSEFCCKGTGYFDGISRMCDVNECQYPRGLGHKNCENTL 293
 QY 190 CTCNPGFTLNEDGRSCODVNECATENPCVQTCVNTYFSFICRCDPGYELEBDYH 248
 Db 294 CSCGVRLSYDVNSCEDINECSS--SPSCQECANITYSGYCYCRQYQALADGHR 352
 QY 249 DECSF--SEFLQHECVNQPGTYFCSCP--PGYILLDDNRSQDINCEHRNHTCNQ 305
 Db 353 DECALPTGHCISYRCINIPESFOCSPSSGYRLAPNRNQCDDIDECVYTGINCINERC 412
 QY 306 YNLGGKCIIDPIRCEEPYLRISDNRC-MCA--ENPGRDQPTLLYRMDVYSGRSVPA 363
 Db 413 FNIQGGKRL--AEPENYRSATRCERLCHENRECSKLPRLTYVHLSPFTIOAPA 471
 QY 364 DIFOMQATTRYPGAYYIFQIKSGNEGREF 391

```

QY 364 DIFOMQATRRYRGAYYIFQIKSGNEGREFYMRQGPISATLVMTRPKPREIOLDLMI 423
DB 602 VIFRMGPSSAVNEDSDSNQALITGNEBGFTRTKVSPHSVVALTKPVPEDL-----LT 656
QY 424 TVNTVINFRG---SSVIRLRIVYS 444
DB 657 TVKMDLSRHGTWSSFFVAKLFIFVS 680

RESULT 6
ID 099K58 PRELIMINARY; PRT: 1174 AA.
AC 099K58;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 21, Last annotation update)
DE Similar to fibulin 2.
GN FBLN2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strusberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC005443; ANH05443.1; -.
DR HSSP: P00736; IAP0.
DR MGD: KGI:95488; EDln2.
DR InterPro: IPR000020; Anaphylatoxin.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR Pfam: PF01821; ANATO; 2.
DR Pfam: PF00008; EGF; 6.
DR SMART: SM00104; ANATO; 3.
DR SMART: SM00179; EGF_CA; 9.
DR SMART: SM00181; EGF; 11.
DR PROSITE: PS01177; ANAPHYLATOXIN_1; 3.
DR PROSITE: PS01178; ANAPHYLATOXIN_2; 3.
DR PROSITE: PS00010; ASX_HYDROXYL; 5.
DR PROSITE: PS01186; EGF_2; 5.
DR PROSITE: PS01187; EGF_CA; 9.
DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
KW SEQUENCE 1174 AA; 126460 MW; 8D628AC710FBA6B8 CRC64;
SQ

Query Match 29.0%; Score 735.5; DB 11; Length 1174;
Best Local Similarity 35.3%; Pred. No. 3.4e-65;
Matches 146; Conservative 61; Mismatches 148; Indels 59; Gaps 11;

```

```

DB 1042 L-RFDCPNVYRVSETKC-----ERTTCODITECQTSPARTHYOLNFQGLLVPAHIFRI 1096
QY 369 QATRRPGAYYIFQIKSGNEGREFYMRQGPISATLVMTRPKPREIOLDLMI 422
DB 1097 GPAPAPAGDTRISLTITGNEGFVTRRLNAYTVGSLSORSVLEPRDPAIDVEM 1150

RESULT 7
ID 073774 PRELIMINARY; PRT: 704 AA.
AC 073774;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE fibulin-1, isoform D.
GN FBLN1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
Gallus.
NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=99120531; Pubmed=9923656;
RA Barth J.L., Argaves K.M., Roark E.F., Little C.D., Argaves W.S.;
RT "Identification of chicken and C. elegans fibulin-1 homologs and
RT characterization of the C. elegans fibulin-1 gene."
RL Matrix Biol. 17:635-646(1998).
DR EMBL: AF051399; AAC05387.1; -.
DR HSSP: P00742; IHCG.
DR InterPro: IPR000020; Anaphylatoxin.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR Pfam: PF01821; ANATO; 2.
DR Pfam: PF00008; EGF; 6.
DR SMART: SM00104; ANATO; 3.
DR SMART: SM00179; EGF_CA; 8.
DR SMART: SM00001; EGF_Like; 1.
DR PROSITE: PS01177; ANAPHYLATOXIN_1; UNKNOWN.1.
DR PROSITE: PS01178; ANAPHYLATOXIN_2; 2.
DR PROSITE: PS00010; ASX_HYDROXYL; 5.
DR PROSITE: PS01186; EGF_2; 3.
DR PROSITE: PS01187; EGF_CA; 8.
DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
KW SEQUENCE 704 AA; 76137 MW; D47D5A3D05E942932 CRC64;
SQ

Query Match 28.5%; Score 722; DB 13; Length 704;
Best Local Similarity 34.6%; Pred. No. 4.4e-64;
Matches 163; Conservative 68; Mismatches 158; Indels 82; Gaps 22;

```

```

OY 56 DMKCVNONGYLCLIPRTNPNYRGYSNPTSPYSGPYPAAPPLISAPNPTISRLICRF 115
DB 68 EMKCIHHYGGYLCCLPNSAAVINDLHG-----EGP-PPVYPPOHNP-----PCBP 111
OY 116 GYOMDESNOCDVDECATDSDHOCNPTQICINTEGGYCSCTDGYWLEGGCOLDIDECRYG 175
DB 112 GYEPDDQSDVDDECAQALHDDRPSDDCHNLSSGYCTCTCDYGRKGPECDVIDECRYR 171
OY 176 YCOQLCANPGYSCTCINPGFTLNEDGRSCODVNECATENPCVOTCVNTYGSFTICRDPG 235
DB 172 YCOHRCVNLPGSFRCCBEPFQGLGNRSCVDVNECDMGAPCEQRCSNXYCTFLCRHQG 231
OY 236 YLEEDGVGSCDMDDECSFSEFLCOHCVNPGTYFECGPPGYILLDNRSQODINEGHR 295
DB 232 YELHRDGFSCSDIDECYSISYLCQYRCVNEGRPSCHCPQGYOLL-ATRLCQDIDECESG 290
OY 296 NHTCNLOOTCYNLOGGFKCIDPICEPEYLRISDNRCMPAENPGCRDOPFTILYRMDV 355
DB 291 AHOCSEAOICVNFHGGYRCVDTNRCVEPTQVSENRCLCPASNLCREOPSIVHRYMTI 350
OY 356 VSGRVPADIFOMQATRRPGATYIPIQISNGEGRFYMQTGPISATLVTPRIKGPPE 415
DB 351 TSERRRPADVFOIQATSVYPGAANAFOIRAGNSQGFYIRQINNVSAMLVLRDPVGPPE 410
OY 416 IQLDLEMITVNTVINFGSSVIRLRIYVSOYPF 448
DB 411 YVLDEMTVNTMSLMSYRASSVLRITVFGAYTF 443

```

RESULT 4

```

OY 0922K8 PRELIMINARY: PRT: 685 AA.
AC 0922K8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Similar to fibulin 1.
GN FBLN1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RA SEQUENCE FROM N.A.
RP Strausberg R.;
RA Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
DB EMBL, BC007140; AA07140.1; -.
DB MGI; MGI:95487; Fbln1.
DR InterPro: IPR000020; Anaphylatoxin.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR001881; EGF-like.
DR Pfam; PF01821; ANATO; 3.
DR Pfam; PF00008; EGF; 6.
DR PROSITE; PS01177; ANAPHYLATOXIN_1; UNKNOWN_3.
DR PROSITE; PS01178; ANAPHYLATOXIN_2; 3.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_4.
DR PROSITE; PS01186; EGF_2; UNKNOWN_3.
DR PROSITE; PS01187; EGF_CA; UNKNOWN_8.
SQ SEQUENCE 685 AA; 75283 MW; EFOD77D7F66B73B8 CRC64;

```

Query Match 29.3%; Score 743; DB 11; Length 685;
 Best Local Similarity 36.1%; Pred. No. 3.2e-66;
 Matches 160; Conservative 65; Mismatches 154; Indels 64; Gaps 17;

```

OY 16 CLSPFGNAQAQCTNGFDLDROSQCCLDIDECRTPEACRGDMCVNONGYLCLIPRTNPNV 75
DB 290 CRP-----KLOCKSGFIOD-ALGNCDIDINECLISAPCPVGOTCINTBESYTC----- 336
OY 76 YRGYSNPNYSPYSGPYPAAPPLISAPNPTISRLICRFQYOM-DESNQCVDVDECATD 134
DB 337 -----QKNVNP-----GGRGYHLNEGTRCVDVDECSPP 365

```

```

OY 135 SHOCNPTQICINTEGGYCSCTDGYWL--LEGOCIDIDEC-RY--GYCOQLCANPGSYS 189
DB 366 AEPCCGHHGCLNLSPGSFRECKAGFYFDGJSRRCVYDINEQOYGRGLCKHKCENTPSSFH 425
OY 190 CTCINPGFTLNEDGRSCODVNECATENPCVOTCVNTYGSFTICRDPGYELEE-DGVHCSDM 248
DB 426 CSCSAGFRLSVGRSCEDVNEC-LNSPCSQECANVGSYQCYCRGYQLSDVDGTCEDI 484
OY 249 DECSF--SEFLCOHCVNPGTYFECSCP-PGYILLDNRSQODINEGHNHTCNLOOTC 305
DB 485 DECALPTGHTICSYRCINIPGSFQCSPPSGYRLAPGRNQCQDIDECVTGIHNCISNETC 544
OY 306 YNLOGGFKCIDPICEPEYLRISDNRC-MCPA-ENPGCRDOPFTILYRMDVVGSRVPA 363
DB 545 FNIGSFRLTS-FECPENYRRSADTCERLPCHENEGCPRLPLRITYHLSFPTNIQVPA 603
OY 364 DIFOMQATRRYGATYIPIQISNGEGRFYMQTGPISATLVTPRIKGPPEIQDLEM- 422
DB 604 VFERMGPPSSAVPGDSMQALITAGNEGFFTRKRVSHHSGVVALTKPIPEPRDILLVYKMD 663
OY 423 -ITVNTVINFGSSVIRLRIYVS 444
DB 664 LYRHGTVSF-----VAKLFIFVS 682

```

RESULT 5

```

OY 08TBH8 PRELIMINARY: PRT: 683 AA.
AC 08TBH8;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Fibulin 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RA SEQUENCE FROM N.A.
RP TISSUE-BRAIN;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL, BC022497; AAH22497.1; -.
SQ SEQUENCE 683 AA; 74423 MW; 2665A3961B6403B4 CRC64;

```

Query Match 29.1%; Score 737; DB 4; Length 683;
 Best Local Similarity 35.4%; Pred. No. 1.3e-65;
 Matches 157; Conservative 67; Mismatches 154; Indels 66; Gaps 17;

```

OY 16 CLSPFGNAQAQCTNGFDLDROSQCCLDIDECRTPEACRGDMCVNONGYLCLIPRTNPNV 75
DB 288 CRP-----KLOCKSGFIOD-ALGNCDIDINECLISAPCPIGHTCINTBESYTC----- 334
OY 76 YRGYSNPNYSPYSGPYPAAPPLISAPNPTISRLICRFQYOM-DESNQCVDVDECATD 134
DB 335 -----QKNVNP-----GGRGYHLNEGTRCVDVDECAPP 363
OY 135 SHOCNPTQICINTEGGYCSCTDGYWL--LEGOCIDIDEC-RY--GYCOQLCANVGSYS 189
DB 364 AEPCCGHHGCLNLSPGSFRECKAGFYFDGJSRRCVYDINEQOYGRGLCKHKCENTPSSFH 425
OY 190 CTCINPGFTLNEDGRSCODVNECATENPCVOTCVNTYGSFTICRDPGYELEE-DGVHCSDM 248
DB 424 CSCSAGFRLSVGRSCEDINECSS-SPCSQECANVGSYQCYCRGYQLSDVDGTCEDI 482
OY 249 DECSF--SEFLCOHCVNPGTYFECSCP-PGYILLDNRSQODINEGHNHTCNLOOTC 305
DB 483 DECALPTGHTICSYRCINIPGSFQCSPPSGYRLAPNNSCQDIDECVTGIHNCISNETC 542
OY 306 YNLOGGFKCIDPICEPEYLRISDNRC-MCPA-ENPGCRDOPFTILYRMDVVGSRVPA 363
DB 543 FNIGGGRCL-AFECPPNYSRNSATRCERLPCHENRECSKPLRLITYHLSFPTNIQAPA 601

```

```

OY 56 DMACVQNGGYLCIPRTNPYRGPNSTPYSGPYAAPLAPNPYPTISRLLICRF 115
DB 68 EMKICINHGGYLCIPRPSAAVINDLHG-----EGP-PPPYPPAQRN-----PCPP 111
OY 116 GYOMDESNOCDVDECATDSHOCNPQICINTEGGYCSCTDGYWLLGECGLDIDECRYG 175
DB 112 GYEPDDQSCVDVDECAQALHDCRPQSDCHNLPGSYCTCPDGYRKIGPECVDDIDECRYR 171
OY 176 YCOQLCANPGSYSCYNPGFTLNEDGRSCQDVNECATENPCVOTCVNTYGSFICRCPG 235
DB 172 YCOHRCVNLPGSFRCCPEPFOLGPNNRSCVDVNECMGAPCEQRCFNSYCTFPLCRHOG 231
OY 236 YELEBEGVHSCDMDCEFSFELCOHECVNPGTYFCSCPPGYILLDDNRSCODINECHR 295
DB 232 YELHNDGFCSDIDECYSYLCQYRCVNPGRFSCPCPGYQLL-ATRLCODIDECRSG 290
OY 296 NHTCNLQOTCYNLQGGFKCIDPIRCEBPYLRISDNRCMCPAENPCGRDQPTILYRDMV 355
DB 291 AHQCSAQTCVNHGRCVDCVTRNCVEPYIQVSENRCICPASNDPLCREQPSIVHRYMTI 350
OY 356 VSGRSVPADIFOMQATTRYGAYIIFQIKSGNEGREFYMRGTGPISATLVMTPLKGPKE 415
DB 351 TBSRSVPADVFOIQATSVYPGAYNAFOIRKNSOGDFYIRQINNVSAVLARPYTGPRE 410
OY 416 IOIDLEMITVNTYINFGSSVIRLRIYVSQYF 448
DB 411 YVLDLEMTVNTMSLMSYRASSVLRITVFGAYTF 443

```

RESULT 2

```

O9J06 PRELIMINARY; PRT; 443 AA.
ID AC O9J06:
AC O9J06:
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE EGF-containing fibulin-like extracellular matrix protein 2.
GN EFEMP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20435063; PubMed=10982184;
RA Katsanis N., Venable S., Smith J.R., Lupski J.R.;
RT "Isolation of a paralog of the Doyme honeycomb retinal dystrophy gene
RT from the multiple retinopathy critical region on 11q13."
EMBL AF109122; AAF65189.1; -.
HUMBL AF109122; AAF65189.1; -.
HSSP: P00736; IAPQ.
DR MGD: MGI:1891209; Efemp2.
DR InterPro: IPR00152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF-Ca.
DR InterPro: IPR001491; Thrombomodulin.
DR Pfam: PF00008; EGF_4.
DR PRINTS: PR00907; THROMBOMODULN.
DR SMART: SM00179; EGF_CA_4.
DR SMART: SM00001; EGF_Like_2.
DR PROSITE: PS00010; ASX_HYDROXYL_4.
DR PROSITE: PS01186; EGF_2_4.
DR PROSITE: PS01187; EGF_CA_6.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation;
KW Matrix protein; Repeat.
SQ SEQUENCE 443 AA; 49452 MW; 5AEC2A91048B336A CRC64;

```

```

Query Match 50.7%; Score 1283; DB 11; Length 443;
Best Local Similarity 49.9%; Pred. No. 6.6e-121;
Matches 226; Conservative 70; Mismatches 135; Indels 22; Gaps 4;
OY 1 MGIKRLVITVITIALCLPSFGNA-----AQTNGFDLDRSGGCLDIDECRTIPEACRG 55

```

```

DB 8 LPGLSLMALFALLLLGASPDDEPDSYECTDGYEMDADSOHCRDVNECLTIPEACKG 67
OY 56 DMACVQNGGYLCIPRTNPYRGPNSTPYSGPYAAPLAPNPYPTISRLLICRF 115
DB 68 EMKICINHGGYLCIPRPSAAVINDLHG-----EGP-PPPYPPAQRN-----PCPP 111
OY 116 GYOMDESNOCDVDECATDSHOCNPQICINTEGGYCSCTDGYWLLGECGLDIDECRYG 175
DB 112 GYEPDDQSCVDVDECAQALHDCRPQSDCHNLPGSYCTCPDGYRKIGPECVDDIDECRYR 171
OY 176 YCOQLCANPGSYSCYNPGFTLNEDGRSCQDVNECATENPCVOTCVNTYGSFICRCPG 235
DB 172 YCOHRCVNLPGSFRCCPEPFOLGPNNRSCVDVNECMGAPCEQRCFNSYCTFPLCRHOG 231
OY 236 YELEBEGVHSCDMDCEFSFELCOHECVNPGTYFCSCPPGYILLDDNRSCODINECHR 295
DB 232 YELHNDGFCSDIDECYSYLCQYRCVNPGRFSCPCPGYQLL-ATRLCODIDECRSG 290
OY 296 NHTCNLQOTCYNLQGGFKCIDPIRCEBPYLRISDNRCMCPAENPCGRDQPTILYRDMV 355
DB 291 AHQCSAQTCVNHGRCVDCVTRNCVEPYIQVSENRCICPASNDPLCREQPSIVHRYMTI 350
OY 356 VSGRSVPADIFOMQATTRYGAYIIFQIKSGNEGREFYMRGTGPISATLVMTPLKGPKE 415
DB 351 TBSRSVPADVFOIQATSVYPGAYNAFOIRKNSOGDFYIRQINNVSAVLARPYTGPRE 410
OY 416 IOIDLEMITVNTYINFGSSVIRLRIYVSQYF 448
DB 411 YVLDLEMTVNTMSLMSYRASSVLRITVFGAYTF 443

```

RESULT 3

```

O9H3D5 PRELIMINARY; PRT; 443 AA.
ID AC O9H3D5:
AC O9H3D5:
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Fibulin-like extracellular matrix protein.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Seibold S., Marx M.;
RT "Cloning of a new fibulin-like gene."
RT Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.
EMBL AF124486; AAC45245.1; -.
HSSP: P35555; IEMN.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF-Ca.
DR InterPro: IPR001491; Thrombomodulin.
DR Pfam: PF00008; EGF_3.
DR PRINTS: PR00907; THROMBOMODULN.
DR SMART: SM00181; EGF_5.
DR SMART: SM00179; EGF_CA_6.
DR SMART: SM00001; EGF_Like_2.
DR PROSITE: PS00010; ASX_HYDROXYL_3.
DR PROSITE: PS01186; EGF_2_4.
DR PROSITE: PS01187; EGF_CA_5.
KW EGF-like domain; Glycoprotein; Hydroxylation; Matrix protein.
SQ SEQUENCE 443 AA; 49535 MW; D91764BF6A8A060 CRC64;

```

```

Query Match 50.0%; Score 1266; DB 4; Length 443;
Best Local Similarity 49.2%; Pred. No. 3.4e-119;
Matches 223; Conservative 74; Mismatches 134; Indels 22; Gaps 5;
OY 1 MGIKRLVITVITIALCLPSFGNA-----AQTNGFDLDRSGGCLDIDECRTIPEACRG 55
DB 8 LPGLSLMALFALLLLGASPDDEPDSYECTDGYEMDADSOHCRDVNECLTIPEACKG 67

```


FT DOMAIN 1335 1379 EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).
FT SITE 847 849 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 304 315 BY SIMILARITY.
FT DISULFID 310 324 BY SIMILARITY.
FT DISULFID 326 339 BY SIMILARITY.
FT DISULFID 550 562 BY SIMILARITY.
FT DISULFID 557 571 BY SIMILARITY.
FT DISULFID 573 586 BY SIMILARITY.
FT DISULFID 592 604 BY SIMILARITY.
FT DISULFID 599 613 BY SIMILARITY.
FT DISULFID 615 628 BY SIMILARITY.
FT DISULFID 634 645 BY SIMILARITY.
FT DISULFID 640 654 BY SIMILARITY.
FT DISULFID 657 669 BY SIMILARITY.
FT DISULFID 675 686 BY SIMILARITY.
FT DISULFID 681 695 BY SIMILARITY.
FT DISULFID 698 709 BY SIMILARITY.
FT DISULFID 715 726 BY SIMILARITY.
FT DISULFID 721 735 BY SIMILARITY.
FT DISULFID 737 750 BY SIMILARITY.
FT DISULFID 756 767 BY SIMILARITY.
FT DISULFID 762 776 BY SIMILARITY.
FT DISULFID 778 791 BY SIMILARITY.
FT DISULFID 797 808 BY SIMILARITY.
FT DISULFID 803 817 BY SIMILARITY.
FT DISULFID 819 832 BY SIMILARITY.
FT DISULFID 838 850 BY SIMILARITY.
FT DISULFID 845 859 BY SIMILARITY.
FT DISULFID 861 873 BY SIMILARITY.
FT DISULFID 879 891 BY SIMILARITY.
FT DISULFID 885 900 BY SIMILARITY.
FT DISULFID 902 915 BY SIMILARITY.
FT DISULFID 921 933 BY SIMILARITY.
FT DISULFID 927 942 BY SIMILARITY.
FT DISULFID 944 957 BY SIMILARITY.
FT DISULFID 963 975 BY SIMILARITY.
FT DISULFID 970 984 BY SIMILARITY.
FT DISULFID 986 1000 BY SIMILARITY.
FT DISULFID 1101 1114 BY SIMILARITY.
FT DISULFID 1109 1123 BY SIMILARITY.
FT DISULFID 1125 1138 BY SIMILARITY.
FT DISULFID 1144 1155 BY SIMILARITY.
FT DISULFID 1150 1164 BY SIMILARITY.
FT DISULFID 1166 1179 BY SIMILARITY.
FT DISULFID 1298 1309 BY SIMILARITY.
FT DISULFID 1304 1318 BY SIMILARITY.
FT DISULFID 1320 1333 BY SIMILARITY.
FT DISULFID 1339 1354 BY SIMILARITY.
FT DISULFID 1349 1363 BY SIMILARITY.
FT DISULFID 1365 1378 BY SIMILARITY.
FT MOD_RES 647 647 HYDROXYLATION.
FT MOD_RES 810 810 HYDROXYLATION.
FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 870 870 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 923 923 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1039 1039 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1394 AA; 152791 MW; DFECA81A0B2C7D1 CRC64; /FTID-CAR_000184.

Query Match 21.5%; Score 544; DB 1; Length 1394;
Best Local Similarity 29.4%; Pred. No. 7e-34;
Matches 128; Conservative 52; Mismatches 134; Indels 122; Gaps 17;

OY 16 CLPBGNAOACTNCFDIDROSGQCLDIDECRTIPACRGDMACVNQNGYLCIPRTNPV 75
DB 562 CINLPVRYTCICYEGRFSEQKCVDICDCTOVQHLG-SQGRCENTEGSFICI----- 614
OY 76 YRGPSNYSYPSYSGPYAAPPLSAPNPJTISRPLI-----CRGQXQ 118
DB 615 -----CPAGFMASEGTNCIDVDECLRPDYCGEGHCVNVTVGAFRCYCDSDGYR 662

OY 119 MDESNOQVDVDECATDSDHOCNPQICINTEGGYTC-SCTDGYMLLEGQCLDIDEC-RYGY 176
DB 663 MTRGRCEIDDECLNS-TC-PDEQCVNPSGYCVCPTBEFRGMNGQCLDIDECLEFN 720
OY 177 CQO-ICANVPGSYSCNPGFTLNEDGRSCODVNECATENPCVQ----- 219
DB 721 CANGDCSNLEGSYMCCHKGYTRTPDHRHCRIDECQOQGNLCVNGQCNTEGSEFRCTGQ 780
OY 220 -----TCVMTYGSFICRCDGYLEEEDGVHCSMDDECSF 253
DB 781 GYLSAKQDCEIDECQHRHLCAHGQCRNTEGSEFCVCDGYPASGLGDHCEIDNECLE 840
OY 254 SEFLCOH-ECVNOPTGYFCSPGYILLDDNRSCODINECHRNHTCMIAQOTCYVLAQGF 312
DB 841 DKSVCQRGDCINTAGSYDCTCPDGF-QLDDNKTQODINECHRP-LCGPQGECLNTEGSEF 898
OY 313 KCI-----DPIRCEEPYLRIS-----DN-----RCMC-----PAENPCR 342
DB 899 HCVCOQGFISADGRCEIDECVNNTVCDSHGFCDMTASGFRCLCYOGFOAPDQGCVCV 958
OY 343 DQPFILYRDMQVYSG 358
DB 959 D-----VNECELLSG 968

Search completed: July 3, 2003, 18:23:14
Job time : 13.3157 secs

```

FT DISULFID 461 474 BY SIMILARITY.
FT DISULFID 480 495 BY SIMILARITY.
FT DISULFID 491 504 BY SIMILARITY.
FT DISULFID 506 519 BY SIMILARITY.
FT DISULFID 525 537 BY SIMILARITY.
FT DISULFID 530 546 BY SIMILARITY.
FT DISULFID 551 565 BY SIMILARITY.
FT VARSPLIC 564 712 BY SIMILARITY.
FT ITLKVSAVNTADPVENEDLOKTTIVGAPNVLPAIRAFI
FT LKGEKRSAYVTLTDSLDGPTVLAQLLRLSKKSKNFT
FT YAAVLIVAAAKRHNTVHPPLMKIR -> QIADYSCKIV
FT CSTEDTECLGNHTRFVLYOFRAVPSIKTISPIEVSRTV
FT MGPEVSVDNLDYGOHRFRIVQERNIGVLCFISGPTV
FT ETIKVNIHTKSRGTGVIILAFNEALIEISVSKYFE (IN
FT ISOFORM D)
SQ SEQUENCE 712 AA: 77009 MM: 52CE8CF8BF296BC5 CRC64:
Query Match 23.7%; Score 599.5; DB 1; Length 712;
Best Local Similarity 27.8%; Pred. No. 2.2e-38;
Matches 155; Conservative 47; Mismatches 153; Indels 203; Gaps 19;
QY 15 LCLBSPG-MAOACTNGFDLDRSGQCLDIDECRIPEACRGMCMVONGCYLIPRTN 73
DB 167 LCHDRGGEKEVCSCRSQFDLPDGMACVYDIDECATLMDCLSEOSQCLVTPGSEFKCI- 222
QY 74 PVYRGPSYNPSTYSGPYPAAPPLASVNPPTISRPILCRGYQMD-ESNOCVYDECA 132
DB 223 -----RTLSGGIYAMDESTERCRDDECN 247
QY 133 TDSHOCNPOTICINTEGGYTC-----SCYDGYWLLEGQCLDIDE 171
DB 248 LGSHDGCPLYOCRCNTQGSYRCDAKCGDGLONPMTGBCTSTICPNGYYPKMGCMNDIDE 307
QY 172 CRGY-----CQO---LCA-----N 183
DB 308 CVYGHNCAGEECVNTPGSFRCQKGNLCAHGYEVNGATGECEDVNECOQVCGSMECTIN 367
QY 184 VPGSYSCNPGFTLNE----- 200
DB 368 LPTGYKKCKGPGYEFNDNAKKRCEDEDECIRKAGHNCDSLAEICNTISSECKCKRGFOLA 427
QY 201 -DGRSCDVNECATE-NECVOTCVNTYGSFTCRCDPGYLEBEGVHCSMDCECF- -S 254
DB 428 SDGRRCEDVNECTTGIAACEQKCAINIPGSIYOCICDRGALDPGKCEBIDICCSIMAGSG 487
QY 255 EFLCOHCVNCPGTYFPCSPRGYILLDNRESCODINEEHNRHNTCNLOOTCVNLGGPRC 314
DB 488 NDLGCMGCIINKGSIYLCQCPGPKIIPGKRTCVYDECA-MGECAGSDKVCVNTLGSFRC 546
QY 315 IDPIRCEEPYLRISDNR-----CMCPAENPCRCRQPPITLYRDMVYVGRSVP- 362
DB 547 -HSIDCTFNTYIHDLNKNRKNRQPSACLPBE--CSNVPLFLTYQFISL-ARAVPIS 600
QY 363 -----ADI--FQMATTTPGAYYIFQIKSGNEGREFFYMRTGP-ISA 402
DB 601 HRAPIITLFKVSAPMAHDEVNFELOKTTIVGAPNVLPAIRAN-----FLQKGEKRSNA 655
QY 403 TLVATRIKGPREFLODL 420
DB 656 VVTLRLSDSPQVTKQL 673

```

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN (1)
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX TISSUE=Fibroblast, and platelet;
RC MEDLINE=90275601; PubMed=2350783;
RA Kanazaki T., Olofsson A., Moren A., Wernstedt C., Hellman U.,
RA Miyazono K., Claesson-Welsh L., Heldin C.-H.;
RT "TGF-beta 1 binding protein: a component of the large latent complex
RL of TGF-beta 1 with multiple repeat sequences.";
CC Cell 61:1051-1061(1990).
CC -1- SUBUNIT: THE LARGE LATENT COMPLEX OF TGF-BETA1 FROM PLATELETS IS
CC COMPOSED OF THE TGF-BETA1 MOLECULE NONCOVALENTLY ASSOCIATED WITH
CC A DISULFIDE-BONDED COMPLEX OF A DIMER OF THE N-TERMINAL PROPEPTIDE
CC OF THE TGF-BETA1 PRECURSOR AND A THIRD COMPONENT DENOTED TGF-
CC BETA1-BP. TGF-BETA1-BP DOES NOT BIND DIRECTLY TO ACTIVE TGF-BETA1.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a short form (shown here) and a
CC long form (AC 014766); are produced by alternative splicing.
CC -1- PTM: CONTAINS HYDROXYLATED ASPARAGINE RESIDUES.
CC -1- PTM: The N-terminus is blocked.
CC -1- SIMILARITY: CONTAINS 16 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M34057; AAA61160.1; -.
DR PIR: A35626; A35626.
DR HSSP: P00750; TPG.
DR GlycoSuiteDB: P22064; -.
DR Genew: HGNC:6714; LTBPL.
DR MIM: 150390; -.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF-Ca.
DR InterPro: IPR002212; Fibril-assoc.
DR Pfam: PF00008; EGF; 15.
DR Pfam: PF00683; TB; 4.
DR SMART: SM00179; EGF_CA; 13.
DR SMART: SM00001; EGF_Like; 4.
DR PROSITE: PS00010; ASX_HYDROXYL; 13.
DR PROSITE: PS00022; EGF_1; 2.
DR PROSITE: PS01186; EGF_2; 11.
DR PROSITE: PS01187; EGF_CA; 15.
KW Growth factor binding; Repeat; EGF-like domain; Hydroxylation; signal;
KW Glycoprotein; Alternative splicing.
FT SIGNAL 20
FT CHAIN 1 1394
FT DOMAIN 300 340
FT REPEAT 348 412
FT DOMAIN 546 587
FT DOMAIN 588 629
FT DOMAIN 630 670
FT DOMAIN 671 710
FT DOMAIN 711 751
FT DOMAIN 752 792
FT DOMAIN 793 833
FT DOMAIN 834 874
FT DOMAIN 875 916
FT DOMAIN 917 958
FT DOMAIN 959 1001
FT REPEAT 1017 1084
FT REPEAT 1097 1139
FT DOMAIN 1190 1262
FT DOMAIN 1140 1180
FT DOMAIN 1294 1334

```

```

RESULT 15
LTBS_HUMAN STANDARD: PRT: 1394 AA.
AC P22064;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Latent transforming growth factor beta binding protein 15 precursor
DE (transforming growth factor beta-1 binding protein 1) (TGF-beta1-Bp-
DE 1).
GN LTBPL.

```

	Matches	161.	Conservative	66.	Mismatches	151.	Indels	93.	Gaps	22.
QY	16	CLPSPGNAQA	CTGFGFDIDROSGOCLDIEDCRTPEACRGMCMVNGGILCIPRTNPV	75						
Db	288	CRP-----	KLOCKSGFIQD-ALGNICIDINEICISAPCPICGHTCINTEGSYTC	334						
QY	76	YRGYSNPNYS	ITPYGGPYAAPPLSAPNPYPLISRLICRFQYM-DESNOCYVDVDECATD	134						
Db	335	-----	QKNVPN-----	CGRGHLNEEGTRCVDVDECAP	363					
QY	135	SHOCNPQIC	INTEGGYTCSTDGWYL-LBGOCLDIEDC-RY--GYCOOLCANPGASYS	189						
Db	364	AEPGKGKRCVNS	SGSFCECKTGTGYEPGISRMCDVNECORPYRPLCGHCNENLGSYL	423						
QY	190	CTCNPGEFLN	DGSRSCODVNECATENPCVQTCVNTGYSFICRCDPYELEE-DGVHCSDM	248						
Db	424	CSCSVGFR	LSYDGRSCDINECCS-SPQSOECANVYGSYGCCRGYGLSDVDGYTCEDI	482						
QY	249	DECSF--SE	LCQHECVNPQPTGYRCSCP-PGYILLDNRSCODINECHRNHTCLQOTC	305						
Db	483	DECALPFG	HCILCSYRCINIPESFQCSPPSGRLAPNRCODIDECYTGHCINSETC	542						
QY	306	YNLAGGFC	EDIPICEEPLYRLISDN-----RCMCPAENPCGRDOPETILYRDMDVY	356						
Db	543	FNIDGARCL	-AFCEPENYRIRSAATLOOEKTDYVNCI-----KSRPDNDYTCVFPVHTI	596						
QY	357	SGRSV-----	PADIFQMA-TTRYPCG-AVYIFQIKSGNREAF-----YMRQYG	398						
Db	597	SHVYISLP	TFREFTRPEEILFLRPAHQANIIDITEGNLRDSFDIIRKYM--DG	654						
QY	399	PISATLV	TRIKGRPREIQDLEH-ITYVNYINFGSSVIRLRIYVSQYF	448						
Db	655	MTVGVRQV	RPIVGPFAVVLKLENNYVGVGVSHR--NVVNVRLTPESEWF	703						

RESULT 14

FBLL_CAEEL STANDARD: PRT: 712 AA.

AC 077469: 077474; Q95N23;
 DT 16-OCR-2001 (Rel. 40, Created)
 DT 16-OCR-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fibulin-1 precursor.
 GN FBUN1 OR F56H11.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 CC Rhabditidae; Peloderae; Caenorhabditis.
 CC NCBI_TaxId=6239;
 RN [1]
 RN SEQUENCE FROM N.A.
 RN MEDLINE=99120531; PubMed=9923656;
 RN Barth J.L., Argaves K.M., Roark E.F., Little C.D., Argaves W.S.,
 RN "Identification of chicken and C. elegans fibulin-1 homologs and
 RN characterization of the C. elegans fibulin-1 gene.";
 RN Matrix Biol. 17:635-646(1998).
 RN [2]
 RN SEQUENCE FROM N.A.
 RN STRAIN=Bristol N2;
 RA Lloyd C.R.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 CC -1 SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX (BY SIMILARITY).
 CC -1 ALTERNATIVE PRODUCTS: 2 ISOFORMS; C (SHOWN HERE) AND D; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1 SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.
 CC -1 SIMILARITY: CONTAINS 9 EGF-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

[illegible]

```

ID      FBIL1_HUMAN      STANDARD.      PRT; 703 AA.
AC      P23142; P23143; P23144; P37888; Q9UGR4;
AD      01-NOV-1991 (Rel. 20, Created)
DT      01-NOV-1991 (Rel. 20, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Fibullin-1 precursor.
GN      Fibullin-1
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX      Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
CX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A. (ISOFORMS A; B AND C).
RX      MEDLINE=91100426; PubMed=2269669;
RA      Argaves W.S., Tran H., Burgess W.H., Dickerson K.;
RT      "Fibullin is an extracellular matrix and plasma glycoprotein with
RL      repeated domain structure."
RN      J. Cell Biol. 111:3155-3164(1990).
RN      [2]
RP      SEQUENCE FROM N.A. (ISOFORM D).
RA      Argaves S.;
RT      Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE OF 567-703 FROM N.A. (ISOFORM C).
RA      Connor R.;
RT      Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN      [4]
RP      SEQUENCE OF 30-44.
RX      MEDLINE=89354537; PubMed=2527614;
RA      Argaves W.S., Dickerson K., Burgess W.H., Ruoslahti E.;
RT      "Fibullin, a novel protein that interacts with the fibronectin
RN      receptor beta subunit cytoplasmic domain."
RN      Cell 58:623-629(1989).
CC      -1- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC      -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; A, B, C AND D (SHOWN HERE); ARE
CC      PRODUCED BY ALTERNATIVE SPLICING. THEY DIFFER ONLY IN THEIR C-
CC      TERMINAL REGIONS.
CC      -1- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.
CC      -1- SIMILARITY: CONTAINS 9 EGF-LIKE DOMAINS.
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X53741; CAAG7770.1; -
DR      EMBL; X53742; CAAG7771.1; -
DR      EMBL; X53743; CAAG7772.1; -
DR      EMBL; U01244; AAB17099.1; -
DR      EMBL; 295331; CAB62960.1; -
DR      PIR; A32826; A32826.
DR      PIR; A36346; A36346.
DR      PIR; B36346; B36346.
DR      PIR; C36346; C36346.
DR      HSSP; P35555; IEMN.
DR      Genew; HGNC:3600; FBIL1.
DR      MIM; 135820;
DR      InterPro; IPR000020; Anaphylatoxin.
DR      InterPro; IPR000152; Asx_hydroxyl.
DR      InterPro; IPR000561; EGF-like.
DR      InterPro; IPR001881; EGF_Ca.
DR      Pfam; PF00008; EGF_6.
DR      Pfam; PF01821; ANATO_3.
DR      SMART; SM00104; ANATO_3.
DR      SMART; SM00179; EGF_CA_7.
DR      SMART; SM00001; EGF_Like_2.
DR      PROSITE; PS00010; ASX_HYDROXYL; 4.
DR      PROSITE; PS00022; EGF_1; FALSE_NEG.
DR      PROSITE; PS01177; ANAPHYLATOXIN_1; 3.
DR      PROSITE; PS01176; ANAPHYLATOXIN_2; 3.

```

[illegible]

NCBI_FaxID-10090;
 [1]
 SEQUENCE FROM N.A.
 STRAIN-MK31;
 MEDLINE-93358897; PubMed-8354280;
 Pan T.-C., Kluge M., Zhang R.Z., Mayer U., Timpl R., Chu M.-L.;
 "Sequence of extracellular mouse protein BM-90/fibulin and its
 calcium-dependent binding to other basement-membrane ligands."
 Eur. J. Biochem. 215:733-740(1993).
 - SUBCELLULAR LOCATION: Secreted; extracellular matrix.
 - ALTERNATIVE PRODUCTS: 4 ISOFORMS: A, B, C AND D (SHOWN HERE); ARE
 PRODUCED BY ALTERNATIVE SPLICING. THEY DIFFER ONLY IN THEIR C-
 TERMINAL REGIONS.
 - SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.
 - SIMILARITY: CONTAINS 9 EGF-LIKE DOMAINS.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

FT DISULFID 250 262 BY SIMILARITY.
 FT DISULFID 268 281 BY SIMILARITY.
 FT DISULFID 275 290 BY SIMILARITY.
 FT DISULFID 296 308 BY SIMILARITY.
 FT DISULFID 314 327 BY SIMILARITY.
 FT DISULFID 321 336 BY SIMILARITY.
 FT DISULFID 343 356 BY SIMILARITY.
 FT DISULFID 362 375 BY SIMILARITY.
 FT DISULFID 369 384 BY SIMILARITY.
 FT DISULFID 386 399 BY SIMILARITY.
 FT DISULFID 405 417 BY SIMILARITY.
 FT DISULFID 413 426 BY SIMILARITY.
 FT DISULFID 428 441 BY SIMILARITY.
 FT DISULFID 447 456 BY SIMILARITY.
 FT DISULFID 452 465 BY SIMILARITY.
 FT DISULFID 467 481 BY SIMILARITY.
 FT DISULFID 487 500 BY SIMILARITY.
 FT DISULFID 496 509 BY SIMILARITY.
 FT DISULFID 511 525 BY SIMILARITY.
 FT DISULFID 531 544 BY SIMILARITY.
 FT DISULFID 538 553 BY SIMILARITY.
 FT DISULFID 558 579 BY SIMILARITY.
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 541 541 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 569 705 PROKTPVRCISCRPDNCAVDPVHTSVTSVLPFR
 EFTPEEPIIFRAVTPYAPNADIIIPITBEGNLDSPDI
 KRYEDGTVGVROVPYIVGPFYATKLEMYVLGVSHR
 NVNVAHIVSEYWF -> RCARLPCHENOCRLDRLITY
 HLPSPNIOVPAVVFVRMGPSSAVPDGSMOLATAGNEGEF
 TTRKVSHTSGVVALTKPIPEPRDLITVYMOLYRGVTVSSF
 VAKLFIYVSAL (IN ISOFORM D).
 SQ SEQUENCE 705 AA; 78056 MM; FD3F06469A4BAE2B C8C64;

Query Match 27.88; Score 703.5; DB 1; Length 705;
 Best Local Similarity 34.98; Pred. No. 2.7e-46;
 Matches 162; Conservative 67; Mismatches 156; Indels 79; Gaps 22;
 16 CLPSPGNAOQNGFPLDRSGOGLDIEDRTPEACRGDMCYNONGYCIPTRTNV 75
 290 CRP-----KLCQSGFIQD-ALGNCIDINEGLSAPFVGOTCTNTEGSTYC----- 336
 76 YRGPSNPSYTPYSGPYPAAPPLSAPNPTISRLICRFGYOM-DESNQCVNDECARD 134
 337 -----QKNVPR-----CGRGYHLINEGTRCVNDECAPP 365
 135 SHQCNPTQICINTEGTYGCTGCTGYWL-LEGCLDIDEC-RY--GYCQQLCANPGSIS 189
 366 AEPCKGHHCLNSPGRCECKAGFYDGLSRCTVDINECORYPGRLCGHKCENTPGSFH 425
 190 CTCNPGFTLNEPDRSCQDVNCAENPCVOTCVMTYGSFICRCQDGYELEE-DGVHCSDM 248
 426 CSCSAGFRLSYDGRSCEDVNEC-LNSPQSCCANVGSYOCYCRGTYQSDVDGYTCEDI 484
 249 DECSF--SEFLCOHECVNOGTYFCSCP-PGYIILDNRSQDINECHRNHTMLOQC 305
 485 DECALPTGHCIGSRCININGSPGSCPSGSRPLAPNRNODDIDECYTGINCINERC 544
 306 YNLOGFKCDLPICEPYLRISDN-----RCM--CAENNGC-RDQFTLLYRDM 353
 545 FNTQGSFRLCS-PECPENYRSADTFRQEKTDYVRCISCRPNDEACVAPHTVSHVYI 603
 354 DVVSGRSV--PADIFOMQATT-RYPG--AYYIFQIKSGNEGREPYM--ROTGPISATLV 405
 604 SLPTFEETPEEETIFRAVTPPLYPANADIIIPITBEGNLDSPDIIRYEGMIVGVYR 663
 406 MTRPIKPREIQLDEM-ITVNTVINRGSSVIRLRYVQYP 448
 664 QVAPYIGPFYAVALKLEMYVLGVVSHR--NVNVAHIVSEYWF 705
 RESULT 13
 FBL_HUMAN

CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX. COMPONENT OF BOTH
 CC BASEMENT MEMBRANES AND OTHER CONNECTIVE TISSUES.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN HEART, PLACENTA AND OVARY.
 CC -1- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 11 EGF-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X82494: CA57876.1; -
 CC HSSP: P00736: 1APQ.
 CC Genew: HGNC:3601; FBLN2.
 CC MIM: 135821; -
 CC InterPro: IPR000020; Anaphylatoxin.
 CC InterPro: IPR000152; Asx_Hydroxyl.
 CC InterPro: IPR000561; EGF-like.
 CC InterPro: IPR001881; EGF_Ca.
 CC Pfam: PF00008; EGF_7.
 CC Pfam: PF01821; ANATO; 2.
 CC SMART: SM00104; ANATO; 2.
 CC SMART: SM00179; EGF_Ca; 3.
 CC SMART: SM00001; EGF_Like; 1.
 CC PROSITE: PS00010; ASX_HYDROXYL; 5.
 CC PROSITE: PS01177; ANAPHYLATOXIN_1; 3.
 CC PROSITE: PS01178; ANAPHYLATOXIN_2; 3.
 CC PROSITE: PS00022; EGF_1; FALSE_NEG.
 CC PROSITE: PS01186; EGF_2; 5.
 CC PROSITE: PS01187; EGF_Ca; 9.
 CC DR Signal: Glycoprotein. Extracellular matrix; Plasma; EGF-like domain;
 CC KW Calcium-binding; Repeat.
 CC FT SIGNAL 1 27 POTENTIAL.
 CC FT CHAIN 28 1184 FIBULIN-2.
 CC FT DOMAIN 28 444 N.
 CC FT DOMAIN 177 444 SUBDOMAIN NA (CYS-RICH).
 CC FT DOMAIN 178 444 SUBDOMAIN NB (CYS-FREE).
 CC FT DOMAIN 445 480 ANAPHYLATOXIN-LIKE 1.
 CC FT DOMAIN 488 519 ANAPHYLATOXIN-LIKE 2.
 CC FT DOMAIN 521 553 ANAPHYLATOXIN-LIKE 3.
 CC FT DOMAIN 604 645 EGF-LIKE 1, CALCIUM-BINDING.
 CC FT DOMAIN 679 718 EGF-LIKE 2.
 CC FT DOMAIN 719 763 EGF-LIKE 3, CALCIUM-BINDING.
 CC FT DOMAIN 764 809 EGF-LIKE 4, CALCIUM-BINDING.
 CC FT DOMAIN 810 857 EGF-LIKE 5, CALCIUM-BINDING.
 CC FT DOMAIN 858 900 EGF-LIKE 6, CALCIUM-BINDING.
 CC FT DOMAIN 901 942 EGF-LIKE 7, CALCIUM-BINDING.
 CC FT DOMAIN 943 981 EGF-LIKE 8, CALCIUM-BINDING.
 CC FT DOMAIN 982 1024 EGF-LIKE 9, CALCIUM-BINDING.
 CC FT DOMAIN 1025 1069 EGF-LIKE 10, CALCIUM-BINDING.
 CC FT DOMAIN 1070 1184 DOMAIN ITI.
 CC FT DISULFID 445 472 BY SIMILARITY.
 CC FT DISULFID 446 479 BY SIMILARITY.
 CC FT DISULFID 459 480 BY SIMILARITY.
 CC FT DISULFID 489 518 BY SIMILARITY.
 CC FT DISULFID 502 519 BY SIMILARITY.
 CC FT DISULFID 521 545 BY SIMILARITY.
 CC FT DISULFID 522 552 BY SIMILARITY.
 CC FT DISULFID 535 553 BY SIMILARITY.
 CC FT DISULFID 608 620 BY SIMILARITY.
 CC FT DISULFID 616 629 BY SIMILARITY.
 CC FT DISULFID 631 644 BY SIMILARITY.
 CC FT DISULFID 683 693 BY SIMILARITY.
 CC FT DISULFID 689 702 BY SIMILARITY.
 CC FT DISULFID 704 717 BY SIMILARITY.
 CC FT DISULFID 723 736 BY SIMILARITY.
 CC FT DISULFID 730 745 BY SIMILARITY.
 CC FT DISULFID 751 762 BY SIMILARITY.
 CC FT DISULFID 768 781 BY SIMILARITY.
 CC FT DISULFID 775 790 BY SIMILARITY.

FT DISULFID 796 808 BY SIMILARITY.
 FT DISULFID 814 827 BY SIMILARITY.
 FT DISULFID 821 836 BY SIMILARITY.
 FT DISULFID 843 856 BY SIMILARITY.
 FT DISULFID 862 875 BY SIMILARITY.
 FT DISULFID 869 884 BY SIMILARITY.
 FT DISULFID 886 899 BY SIMILARITY.
 FT DISULFID 905 917 BY SIMILARITY.
 FT DISULFID 913 926 BY SIMILARITY.
 FT DISULFID 928 941 BY SIMILARITY.
 FT DISULFID 947 956 BY SIMILARITY.
 FT DISULFID 952 965 BY SIMILARITY.
 FT DISULFID 967 980 BY SIMILARITY.
 FT DISULFID 986 998 BY SIMILARITY.
 FT DISULFID 994 1007 BY SIMILARITY.
 FT DISULFID 1009 1023 BY SIMILARITY.
 FT DISULFID 1029 1042 BY SIMILARITY.
 FT DISULFID 1036 1051 BY SIMILARITY.
 FT DISULFID 1056 1068 BY SIMILARITY.
 FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 507 507 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1035 1035 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 1184 AA; 126543 MW; CA48490A53F9ECSD CRC64;
 Query Match 28.0%; Score 709.5; DB 1; Length 1184;
 Best Local Similarity 34.3%; Pred. No. 1,6e-46;
 Matches 146; Conservative 61; Mismatches 156; Indels 63; Gaps 12;
 QY 15 ICLPSRGN---AQAQCTNGFDLDRSGGCLDDECRITPEACRGDMCYNGYICIP 70
 DB LCQNTKGSFCCQAKORCMDF-LQDEGNCVVDINETSLEPCRPSPFCINTVGSYTC-- 836
 QY 71 RTNPPYRGPSNPSTPYSGPYPAAPPLSAPNPTISPLICRFQYQ-MDESGCVDVD 129
 DB 837 -----QRNPPLICARITHASDQAKCVDN 860
 QY 130 ECATSDHOCNPQTICINTEGCTCSDGYW--LEGGCLDDECRYG--YCQOLCANY 184
 DB 861 ECETGVHNCGEQYOVHNLPGSYRCDCAKGFORDAFGRGCIDVNECWSAQRGLCQHCEMT 920
 QY 185 PGYSCTCNPQFTLNEDGRSCODVNECATENPCYQCVNTYSFICRCRPGYLEDDGYH 244
 DB 921 LGSYRCSASAGFLAADGRKCEDVNECEQR--CSQECANIVSYCYCRQYOLADGHT 979
 QY 245 CSDMECS-FSEFLQHCVCVNPQGYFSCSP-PGYTLDDNNSCODINECHRNHTCNQ 302
 DB 980 CTIDDECAQAGAILCTFRCLANPGSYOCAPQGYTMANGSCVDVDCALGTHNCSEA 1039
 QY 303 QTCYNLOGGFKCIDIRCEEPYLRISDNRCMKPAENPGCRD-----QPTILYRDMVY 356
 DB 1040 ETCNHIQSEFKCL-RFECPPNIVQVSKTKC---ERTTCHDFLECONSPPARTHTYQLNFQ 1094
 QY 357 SGRSVPADIFQMOATTRYGAYIYIPIQKSGNEGRFYMROTGISATLYMTPIKGPRT 416
 DB 1095 TGLVPAHIFRIGAPAPFTGDTLNIITKGNDEGIFGRRLNATYGVYLOAVLEPRDF 1154
 QY 417 QLDDEM 422
 DB 1155 ALDDEM 1160
 RESULT 12
 FBL_MOUSE
 ID FBL_MOUSE STANDARD; PRT; 705 AA.
 AC Q08879; Q08879; 008879 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fibrin-1 precursor (Basement-membrane protein 90) (BM-90).
 GN FBLN1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.


```

FT DOMAIN 400 441 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 442 481 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 482 525 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 526 570 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
FT DISULFID 33 59 BY SIMILARITY.
FT DISULFID 34 66 BY SIMILARITY.
FT DISULFID 47 67 BY SIMILARITY.
FT DISULFID 76 107 BY SIMILARITY.
FT DISULFID 89 108 BY SIMILARITY.
FT DISULFID 110 134 BY SIMILARITY.
FT DISULFID 111 141 BY SIMILARITY.
FT DISULFID 124 142 BY SIMILARITY.
FT DISULFID 181 191 BY SIMILARITY.
FT DISULFID 187 200 BY SIMILARITY.
FT DISULFID 202 215 BY SIMILARITY.
FT DISULFID 221 234 BY SIMILARITY.
FT DISULFID 228 243 BY SIMILARITY.
FT DISULFID 249 261 BY SIMILARITY.
FT DISULFID 267 280 BY SIMILARITY.
FT DISULFID 274 289 BY SIMILARITY.
FT DISULFID 295 307 BY SIMILARITY.
FT DISULFID 313 326 BY SIMILARITY.
FT DISULFID 320 335 BY SIMILARITY.
FT DISULFID 342 355 BY SIMILARITY.
FT DISULFID 361 374 BY SIMILARITY.
FT DISULFID 368 383 BY SIMILARITY.
FT DISULFID 385 398 BY SIMILARITY.
FT DISULFID 404 416 BY SIMILARITY.
FT DISULFID 412 425 BY SIMILARITY.
FT DISULFID 427 440 BY SIMILARITY.
FT DISULFID 446 455 BY SIMILARITY.
FT DISULFID 451 464 BY SIMILARITY.
FT DISULFID 466 480 BY SIMILARITY.
FT DISULFID 486 499 BY SIMILARITY.
FT DISULFID 495 508 BY SIMILARITY.
FT DISULFID 510 524 BY SIMILARITY.
FT DISULFID 530 543 BY SIMILARITY.
FT DISULFID 537 552 BY SIMILARITY.
FT DISULFID 557 569 BY SIMILARITY.
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 536 536 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 540 540 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 638 638 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 684 AA: 75623 MW: 1638D7A094739199 CRC64;

```

Query Match 30.7%; Score 778.5; DB 1; Length 684;
 Best Local Similarity 35.9%; Pred. No. 5.3e-52;
 Matches 162; Conservative 66; Mismatches 154; Indels 69; Gaps 16;

```

QY 15 LCLSPGN---AQAQCTNGFDLDRSGQGLDIDECRTPEACRGDMCVNONGYLCLIP 70
DB 279 ICONPGSFRCRPRKLCQMGFIOD-ALGNCIDINECLSTNMPCPAQDICTNGSYTC-Q 336
QY 71 RTNPPYRBPYSNMPYTPYSGPPAPAPRISAPRYPIISRLICRFGYQDDE-SNOCYVDV 129
DB 337 RISP-----SCGRGYHLNEDTRCVDV 359
QY 130 ECATDSHOCNPTQICINTEGGYTCSDTGYM--LLEGQCLDIDECRY---GYCQCLCANV 164
DB 360 ECSSSDQCGGSHVCINGNPGNRCKCKSGYSPFDVISRITDIDINECRRYRGLCAHKCENT 419
QY 185 PGSTSCITCNPGFTLNEDGRSCDVNCAENPCVOTCVNTYSGFICRCDPGYELER-DGV 243
DB 420 PGSYCTCTMGFKLSIDRSRCDLNECES-SPCSQDCANVYSGYQCYCRGFLSDIDGT 478
QY 244 HCSMDDECSF--SEFLQHECVNPGTYFCSCP-FGYILLDDNRSCQDINECHRNHTCN 300
DB 479 SCEPIDICALPTGHCISFRCTININGSPQCTPSTGYRLAPAPARNQDIDECVAETHNCS 538
QY 301 LQQTCTYNGGFKCIDPICEEPPYLRISDNRC-MCPA-ENPGCRDPFTILVRDMVYSG 358
DB 539 FNEICFNQGGFRCLIS-LECEPBNYKSGDTRCERLPCNENKRCQSLPRLITYHLSFPIN 597

```

```

QY 359 RSPVADIFQOATRRIPGAVYIFQIKSGNREGRYMQGPISATLVWTRPKGPRIQL 418
DB 598 IQVPTDIFRPGSNAPVGDKILLISISGNQEGFTTKVNHSGIYVWQRDITPRDLL 657
QY 419 DLEMI-----TVNTVINFRRSSVIRLRIYYS 444
DB 658 TIOMQLTRHGIVNTFTI-----AKLFEVYS 681

RESULT 10
FBL2_MOUSE
ID FBL2_MOUSE STANDARD; PRT; 1221 AA.
AC P37889; Q9WU12.
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fbln2-2 precursor.
GN FBLN2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-35.
RC TISSUE=Fibroblast.
RX MEDLINE=94064787; PubMed=8245130;
RA Pan T.-C., Sasaki T., Zhang R.-Z., Faessler R., Timpl R., Chu M.-L.;
RT "Structure and expression of fibulin-2, a novel extracellular matrix
RT protein with multiple EGF-like repeats and consensus motifs for
RT calcium binding."
RL J. Cell Biol. 123:1269-1277(1993).
RN [2]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=99373786; PubMed=10406956;
RA Graessel S., Sicot F.-X., Gotta S., Chu M.-L.;
RT "Mouse fibulin-2 gene. Complete exon-intron organization and promoter
RT characterization."
RL Eur. J. Biochem. 263:471-477(1999).
CC -1- FUNCTION: ITS BINDING TO FIBRONECTIN AND SOME OTHER LIGANDS IS
CC CALCULON DEPENDENT.
CC -1- SUBUNIT: HOMOTRIMER; DISULFIDE-LINKED.
CC -1- SUBCELLULAR LOCATION: Extracellular matrix.
CC -1- ALTERNATIVE PRODUCTS: At least 2 isoforms: 1 (shown here) and
CC 2/EGF-less: are produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: COMPONENT OF BOTH BASEMENT MEMBRANES AND OTHER
CC CONNECTIVE TISSUES.
CC -1- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 11 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X75285; CAA53040.1; -
DR EMBL; AF135253; AAD34456.1; -
DR EMBL; AF135253; AAD34456.1; JOINED.
DR EMBL; AF135240; AAD34456.1; JOINED.
DR EMBL; AF135241; AAD34456.1; JOINED.
DR EMBL; AF135242; AAD34456.1; JOINED.
DR EMBL; AF135243; AAD34456.1; JOINED.
DR EMBL; AF135244; AAD34456.1; JOINED.
DR EMBL; AF135245; AAD34456.1; JOINED.
DR EMBL; AF135247; AAD34456.1; JOINED.
DR EMBL; AF135248; AAD34456.1; JOINED.
DR EMBL; AF135249; AAD34456.1; JOINED.
DR EMBL; AF135250; AAD34456.1; JOINED.
DR EMBL; AF135251; AAD34456.1; JOINED.
DR EMBL; AF135252; AAD34456.1; JOINED.

```



```

DR PROSITE: PS00010; ASX-HYDROXYL; 4.
DR PROSITE: PS00022; EGF_1; FALSE_NEG.
DR PROSITE: PS01186; EGF_2; 4.
DR PROSITE: PS01187; EGF_CA; 6.
DR Repeat; EGF-like domain; Calcium-binding; Glycoprotein; Signal.
FT SIGNAL 1 17
FT CHAIN 18 493
FT DOMAIN 26 71
FT DOMAIN 173 213
FT DOMAIN 214 253
FT DOMAIN 254 293
FT DOMAIN 294 333
FT DOMAIN 334 378
FT DISULFID 177 190
FT DISULFID 184 199
FT DISULFID 201 212
FT DISULFID 218 228
FT DISULFID 224 237
FT DISULFID 239 252
FT DISULFID 258 268
FT DISULFID 264 277
FT DISULFID 278 292
FT DISULFID 305 318
FT DISULFID 320 332
FT DISULFID 338 350
FT DISULFID 344 359
FT DISULFID 365 377
SQ SEQUENCE 493 AA; 54596 MW; 22DAFD70BACFCA5 CRC64;

Query Match 43.8%; Score 1109.5; DB 1; Length 493;
Best Local Similarity 42.3%; Pred. No. 2,7e-77;
Matches 206; Conservative 73; Mismatches 161; Indels 47; Gaps 5;

7 ILTVLLALCLPSPNAQACCTNGEDLDQSCGCLDIDECRTPEACRGDMCMVONGY 66
9 MLLTALVLSQVETITTYTQCTDGYEMDPVROCKDIDCDIDPCKGKMKCVNHNGY 68
67 LCIPRTNVPYRPNPSTPYSPYSPYSPYSPYSPYSPYSPYSPYSPYSPYSPYSPY 102
69 LCLPRTAQIIYNNPQOETPRAEASSGATGTIARSMATSGVITGGGFIASATVAGP 128
103 NYPT-----ISRLICRFGYQMDSESNOCVDVDECATDSHOCNPTQ 142
129 EVQGRNMFVIRNPADPQRIIPSNPSHRIQCAAGYEGSHVNCQDIDECTSGHNRLQ 188
143 ICINTEGGYTCSDGYWLLBEGCCDIDECRT-GYCOQLCANVPYSCTCNPFTLNED 201
189 VCIILRGSTFCHCLPQYKRGBOQVDECSVPPYCHQGCVTMPGSGFYCQCNPGFLAAN 248
202 GRSCQDVNECATENPCVOTCVNTYGSFICRCPGYLEEDGYHCSMDSECFSEFLCOH 261
249 NYTCVDINECDASNOCAQCYNIIGSFICQCNQOGLSELDRLNCEIDIDECRTSSYLCQ 308
262 CVNPGGYFFCCSPGYILLDDNSQDINECHRNHTCNLQOCTYNLAGCFICDIBCE 321
309 CVNPGGYFFCCSPGYILLDDNSQDINECHRNHTCNLQOCTYNLAGCFICDIBCE 366
322 EPLRLIDNRMCAPENPCRDOPFTILYRDMDVYSGRSPADIDOMOTVTPGAYTF 381
367 DPVLYLTSENKCVSVNTMCRDPOSIVYKYNINISDSVPSDIDQIQTITTYANTINF 426
382 QISGNGREFFYMQGTISATLVNTRPIKGPRIQDLEMTITVNTVNFGRSSVLRNI 441
427 RISGNGREFFYMQGTISATLVNTRPIKGPRIQDLEMTITVNTVNFGRSSVLRNI 486
442 YVSGYFF 448
487 IVGPFF 493

```

RESULT 8

```

FBL3_HUMAN
ID FBL3_HUMAN STANDARD; PRT; 493 AA.
AC Q12805;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE EGF-containing fibulin-like extracellular matrix protein 1 precursor
DE (Fibulin-3) (FBLN-3) (Fibillin-like protein) (Extracellular protein
DE S1-5).
GN EFEMP1 OR FBLN3 OR FBNL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI-TaxId=9606;
[1]
RP SEQUENCE FROM N.A., AND POSSIBLE ALTERNATIVE SPLICING.
RC TISSUE-Skin;
RX MEDLINE=95097983; PubMed=7799918;
RA Lecka-Czerwik B., Lumpkin C.K. Jr., Goldstein S.;
RT "An overexpressed gene transcript in senescent and quiescent human
RT fibroblasts encoding a novel protein in the epidermal growth factor-
RT like repeat family stimulates DNA synthesis."
RL Mol. Cell. Biol. 15:120-128(1995).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97001163; PubMed=8812496;
RA Ikegawa S., Toda T., Okui K., Nakamura Y.;
RT "Structure and chromosomal assignment of the human S1-5 gene (FBNL)
RT that is highly homologous to fibrillin."
RL Genomics 35:590-592(1996).
[3]
RN RN
RP SEQUENCE FROM N.A.
RX MEDLINE=20068041; PubMed=10601734;
RA Gittay R., Timpl R., Kostka G.;
RT "Sequence, recombinant expression and tissue localization of two novel
RT extracellular matrix proteins, fibulin-3 and fibulin-4."
RL Matrix Biol. 18:469-480(1999).
[4]
RP VARIANT DHRD/MLYT TRP-345, AND VARIANT PHE-220.
RX MEDLINE=99295941; PubMed=10369267;
RA Stone E.M., Lotery A.J., Munier F.L., Heon E., Pignat B., Guymer R.H.,
RA Vandenberg K., Cousin P., Nishimura D., Swiderski R.E., Silvestri G.,
RA Mackey D.A., Hageman G.S., Bird A.C., Sheffield V.C.,
RA Schorderet D.F.;
RT "A single EFEMP1 mutation associated with both malattia Leventinese
RT and Doyme honeycomb retinal dystrophy."
RL Nat. Genet. 22:199-202(1999).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- ALTERNATIVE PRODUCTS: 4 isoforms; 1 (shown here), 2, 3 and 4; may
CC be produced by alternative splicing.
CC -1- DISEASE: DEFECTS IN EFEMP1 ARE A CAUSE OF DOYNE HONEYCOMB RETINAL
CC DYSTROPHY (DHRD) ALSO KNOWN AS MALATTIA LEVENTINESE (MLYT OR ML),
CC AN AUTOSOMAL DOMINANT DISEASE CHARACTERIZED BY YELLOW-WHITE
CC DEPOSITS KNOWN AS DRUSEN THAT ACCUMULATE BENEATH THE RETINAL
CC PIGMENT EPITHELIUM.
CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U03877; AAA65590.1; --
DR HSSP: P35555; 1EMN.
DR Genew: HGNC:3218; EFEMP1.
DR MIM: 601548; --
DR MIM: 126600; --
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.

```


16-OCT-2001 (Rel. 40, Last annotation update)
 EGF-containing fibulin-like extracellular matrix protein 2 precursor
 (Fibulin-4) (FIBL-4) (UPH1 protein).
 EFEMP2 OR FBLN4.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Melanoma;
 RX MEDLINE=20068041; PubMed=10601734;
 RA Gilbey R., Timpi R., Koska G.;
 RT "Sequence, recombinant expression and tissue localization of two novel
 extracellular matrix proteins, fibulin-3 and fibulin-4.";
 RL Matrix Biol. 18:469-480(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC Zemel R., Shaul Y.;
 RX Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20435063; PubMed=10982184;
 RA Katsanis N., Venable S., Smith J.R., Lupski J.R.;
 RT "Isolation of a paralog of the Doyme honeycomb retinal dystrophy gene
 from the multiple retinopathy critical region on 11q13.";
 RL Hum. Genet. 106:66-72(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX Strausberg R.;
 RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RL -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.sdb-std.ch/announce/>
 or send an email to license@sdb-std.ch).
 CC -----
 CC EMBL: AJ132819; CA10791.2; -
 DR EMBL: AF093119; AAC62108.1; -
 DR EMBL: AF109121; AAF65188.1; -
 DR EMBL: BC010456; AAH10456.1; -
 DR HSSP: P35555; 1EMN.
 DR GeneW: HGNC:3219; EFEMP2.
 DR MIM: 604633; -
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR001491; Thrombospondin.
 DR Pfam: PF00008; EGF; 4.
 DR PRINTS: PR00907; THROMBOSPDNL.
 DR SMART: SM00179; EGF_CA; 4.
 DR SMART: SM00001; EGF_Like; 2.
 DR PROSITE: PS00010; ASX_HYDROXYL; 4.
 DR PROSITE: PS00022; EGF_1; FALSE_NEG.
 DR PROSITE: PS01186; EGF_2; 4.
 DR PROSITE: PS01187; EGF_CA; 6.
 KW Repeat; EGF-like domain; Calcium-binding; Glycoprotein; signal.
 FT SIGNAL 1 25
 FT CHAIN 26 443
 FT EGF-CONTAINING FIBULIN-LIKE EXTRACELLULAR
 FT MATRIX PROTEIN 2.
 FT EGF-LIKE 1, DIVERGENT.
 FT EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
 FT EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
 FT EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
 FT EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
 FT EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
 FT BY SIMILARITY.
 FT DISULFID 127 140

FT DISULFID 134 149 BY SIMILARITY.
 FT DISULFID 151 162 BY SIMILARITY.
 FT DISULFID 168 177 BY SIMILARITY.
 FT DISULFID 173 186 BY SIMILARITY.
 FT DISULFID 188 201 BY SIMILARITY.
 FT DISULFID 207 217 BY SIMILARITY.
 FT DISULFID 213 226 BY SIMILARITY.
 FT DISULFID 228 241 BY SIMILARITY.
 FT DISULFID 247 258 BY SIMILARITY.
 FT DISULFID 254 267 BY SIMILARITY.
 FT DISULFID 269 281 BY SIMILARITY.
 FT DISULFID 287 300 BY SIMILARITY.
 FT DISULFID 294 309 BY SIMILARITY.
 FT DISULFID 315 327 BY SIMILARITY.
 FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 394 394 A -> T (IN REF. 1).
 FT CONFLICT 5 5 EMPDSDSH -> TOTAN (IN REF. 2).
 FT CONFLICT 103 111 AOHDPNCP -> VNIQPLPT (IN REF. 2).
 FT CONFLICT 294 294 C -> W (IN REF. 2).
 FT CONFLICT 354 356 RSV -> AER (IN REF. 2).
 FT CONFLICT 355 355 S -> R (IN REF. 3).
 SQ SEQUENCE 443 AA; 49391 MW; 9E9AC2393780D3B8 CRC64;
 Query Match 50.9%; Score 1289; DB 1; Length 443;
 Best Local Similarity 49.9%; Pred. No. 5.7e-91;
 Matches 226; Conservative 74; Mismatches 131; Indels 22; Gaps 5;
 QY 1 MPGIRILVTTLALCLPSPGNAQ-----AQTNGEPDLDRSGGCLDIDECRTPEACRG 55
 DB 8 LFGSLLMLALLLLLSASBPDSSEPDSTYECIDGYEMWDSQHCRCRVNCELTIPKCKG 67
 QY 56 DMACVNQNGCYICIPPTNPNVYKGPYSPYSGPYAPAAPLADPNPTISPLICRF 115
 DB 68 EMKCIHNYGYCLCPRSAVAVINDLHG-----EGP-PPVPPAQNHN-----PCPP 111
 QY 116 GRPMESNOCVAVDECATDSHOCPNPOICINTEGGTSCSTDGYWLLGSCDIDECRYG 175
 DB 112 GTEPDDDSQVYDDECAQLHDCRPSQDCHNLPGSTQCTPDDGRKIGPCVDDECRYR 171
 QY 176 YCOQLCANYPGSGYSCINPQFTLNEDGRSCQDVNTEATNPQVOTCVNTGSGFTCRDGP 235
 DB 172 YCOHRCVNLPGSFRCCGCEPQFQGLGPNRSCVDVNECDMGAPCGRQCNSTGTFLCRHQG 231
 QY 236 YLEDEDVHGSCMDESESEFCLQHCVCVNPQGYFSCPPGYILLDDNSCODINEGHR 295
 DB 232 YELHRDGFSCSDIDECSSSYLQYRCVNEPGRFSCHCPQGYQL-ATRCODIDEECSG 290
 QY 296 NHTCNLOQTCVNLQGGFKCIDPIRCEPFLRISDNRCMCPAENPGCRDOPFTILYRMDV 355
 DB 291 AHQCEAQTCAVPHFGYRCVDINRCYEPITQVSENKCLCPASNPCLCREQSSIVHRYMTI 350
 QY 356 VSGRSVPADIFQMATTRYGAYIYFIQKSGNREGRFYMRQTPISATLVMTPIKPRE 415
 DB 351 TSERSVPADVFOIQATSYVPGAYVNAFOIRANGSGQGFYIHOINNVASMLVIAPVTPRE 410
 QY 416 IQLDLEMTNVNVINPGRSSVIRLRYVQYPR 448
 DB 411 YVLDLEMTNLSMYSRASSVIRLTVFGAYTP 443
 RESULT 6
 FBLA_MOUSE STANDARD: PRT; 443 AA.
 AC Q9WVJ9;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE EGF-containing fibulin-like extracellular matrix protein 2 precursor
 (Fibulin-4) (FIBL-4) (Mutant p53 binding protein 1).
 GN EFEMP2 OR FBLN4 OR MBP1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

DB 1 MPEIKRLITVTLALMLPHPNNAOQCTNGFDLDRSGQCLDIDECRTIPEACRGDMKCV 60
 QY 61 NONGGLICPRTNPNVGRGPNSTPYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
 DB 61 NONGGLICPRTNPNVGRGPNSTPYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
 QY 121 ESNOCVDEVCANDSHQCNPTQICINTEGGYCTCTDQWYLLLEGQCLDIDECRYGCOOL 180
 DB 121 EGNQCVDEVCANDSHQCNPTQICINTEGGYCTCTDQWYLLLEGQCLDIDECRYGCOOL 180
 QY 181 CANPVSSTCTCPNGFTLNEDGRSCQDVNCEATENPCVQTCVNTYGSFTICRDPGELEE 240
 DB 181 CANPVSSTCTCPNGFTLNEDGRSCQDVNCEATENPCVQTCVNTYGSFTICRDPGELEE 240
 QY 241 DGVHCSDMDECSSEFLCOHECVNQPSTYSCPGYLLDNDNSCODINCEHNHCN 300
 DB 241 DGVHCSDMDECSSEFLCOHECVNQPSTYSCPGYLLDNDNSCODINCEHNHCN 300
 QY 301 LQOTCYNLGGFKCIDPICEEPLRISDNRCMCPAENPCGRDQPFITLYRDMVYSGRS 360
 DB 301 LQOTCYNLGGFKCIDPICEEPLRISDNRCMCPAENPCGRDQPFITLYRDMVYSGRS 360
 QY 361 VPADIFOMATTPYPCGATYTFQIKSGNGBEFYMRQGTGPISATLYMTRPIKPREIQLDL 420
 DB 361 VPADIFOMATTPYPCGATYTFQIKSGNGBEFYMRQGTGPISATLYMTRPIKPREIQLDL 420
 QY 421 EMITVNTVINFGRSSVRLRIYVSQYPF 448
 DB 421 EMITVNTVINFGRSSVRLRIYVSQYPF 448

RESULT 4
 FBIL CRIGR STANDARD: PRT: 443 AA.

AC 055058;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE EGF-containing fibulin-like extracellular matrix protein 2 precursor
 DE (Fibulin-4) (FBLN-4) (H411 protein).
 GN EFEMP2 OR FBLN4.
 OS Cricetus griseus (Chinese hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Cricetulus.
 OC NCBI_TaxID=10029;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 A Heline H., Delude R.L., Monks B., Golenbock D.T.;
 A submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AF046870; AAC03101.1; .
 DR HSSP: P00736; IAP0.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000361; EGF-like.
 DR InterPro: IPR001881; EGF_Ca.
 DR Pfam: PF00008; EGF_4.
 DR SMART: SM00179; EGF_CA; 4.
 DR SMART: SM00001; EGF_Like; 2.
 DR PROSITE: PS00010; ASX_HYDROXYL; 4.
 DR PROSITE: PS00022; EGF_1; FALSE_NEG.
 DR PROSITE: PS01186; EGF_2; 4.

DR PROSITE: PS01187; EGF_CA; 6.
 KW Repeat; EGF-like domain; Calcium-binding; Glycoprotein; Signal.
 FT SIGNAL 1 25
 FT CHAIN 26 443
 FT
 FT DOMAIN 36 81
 FT DOMAIN 123 163
 FT DOMAIN 164 202
 FT DOMAIN 203 242
 FT DOMAIN 243 282
 FT DOMAIN 283 328
 FT DOMAIN 329 368
 FT DISULFID 127 140
 FT DISULFID 134 149
 FT DISULFID 151 162
 FT DISULFID 168 177
 FT DISULFID 173 186
 FT DISULFID 188 201
 FT DISULFID 207 217
 FT DISULFID 213 226
 FT DISULFID 228 241
 FT DISULFID 247 258
 FT DISULFID 254 267
 FT DISULFID 269 281
 FT DISULFID 287 300
 FT DISULFID 294 309
 FT DISULFID 315 327
 FT CARBOHYD 198 198
 FT CARBOHYD 394 394
 SQ SEQUENCE 443 AA; 49432 MW; 0B0FE5D7323D9E5F CR64;
 Query Match 50.9%; Score 1290; DB 1; Length 443;
 Best Local Similarity 50.1%; Pred. No. 4.8e-91;
 Matches 227; Conservative 71; Mismatches 133; Indels 22; Gaps 5;

QY 1 MPEIKRLITVTLALMLPHPNNAOQCTNGFDLDRSGQCLDIDECRTIPEACRG 55
 DB 8 LPSLLMALMLLILGLAASPDQSEPDSTCTDGEYMDADSOHCRDNECLTIEACRG 67
 QY 56 DMKCVNONGGLICPRTNPNVGRGPNSTPYSGPYPAAPPLSAPNPTISRPLICRF 115
 DB 68 EMKCNHGYGGLICLPRSAVYNDLHG-----EGP-PPVPPAQNPN-----PCRP 111
 QY 116 GYOMDESNOGVDECATDHSQCNPTQICINTEGGYCTCTDQWYLLLEGQCLDIDECRYG 175
 DB 112 GYEPDEGSCVDVDECAQALHDCRPSQDCHMLPGSYCTCPDGYRKVGECVIDECRYR 171
 QY 176 YCOQLCANVPGSSTCTNPGFTLNEDGRSCQDVNCEATENPCVQTCVNTYGSFTICRDPG 235
 DB 172 YCOHRCVNLGFSFRCQCEPFGOLGPNPNSCYDVNCEATENPCVQTCVNTYGSFTICRDPG 231
 QY 236 YELLEDGVHCSDMDECSSEFLCOHECVNQPSTYSCPGYLLDNDNSCODINCEHR 295
 DB 232 YELHRDVSFSDIDECYSTYLCOYRCVNEGRSCHKPQGYQL-ATRLCODIDECETG 290
 QY 296 NHTCNLQOTCYNLGGFKCIDPICEEPLRISDNRCMCPAENPCGRDQPFITLYRDMV 355
 DB 291 AHQCSQAQTCVNFHGRVDTNRCVPEYQVSDMRCPYSNPLCRQEPSSIVHRYSI 350
 QY 356 VSGRSVPADIFOMATTPYPCGATYTFQIKSGNGBEFYMRQGTGPISATLYMTRPIKPRE 415
 DB 351 TSEKSVADVFQIOLATSVYPCGATYTFQIKSGNGBEFYMRQGTGPISATLYMTRPIKPRE 410
 QY 416 IOLDLEMITVNTVINFGRSSVRLRIYVSQYPF 448
 DB 411 YVLDLEMITVNTVINFGRSSVRLRIYVSQYPF 443

RESULT 5
 FBIL HUMAN STANDARD: PRT: 443 AA.
 ID PBL4_HUMAN
 AC 095967; 075967;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)

```

FT DOMAIN 247 287 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT SITE 288 333 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT DISULFID 131 144 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 138 143 BY SIMILARITY.
FT DISULFID 155 166 BY SIMILARITY.
FT DISULFID 172 181 BY SIMILARITY.
FT DISULFID 177 190 BY SIMILARITY.
FT DISULFID 192 205 BY SIMILARITY.
FT DISULFID 211 221 BY SIMILARITY.
FT DISULFID 217 230 BY SIMILARITY.
FT DISULFID 232 245 BY SIMILARITY.
FT DISULFID 251 262 BY SIMILARITY.
FT DISULFID 258 271 BY SIMILARITY.
FT DISULFID 273 286 BY SIMILARITY.
FT DISULFID 292 305 BY SIMILARITY.
FT DISULFID 299 314 BY SIMILARITY.
FT DISULFID 320 332 BY SIMILARITY.
FT CARBOHYD 283 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 238 238 L -> P (IN REF. 2).
SQ SEQUENCE 448 AA; 50160 MM; E6BC68F7B14B714 CRC64;

Query Match 95.0%; Score 2406; DB 1; Length 448;
Best Local Similarity 94.2%; Pred. No. 7.8e-176;
Matches 422; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

OY 1 MFGIRLITVLTIALCLPSPGNNAOCTNFGDLDRSGOGLDIDECRTIPACGDMNCV 60
DB 1 MFGRLITVLTIALMLPSPGNNAOCTNFGDLDRSGOGLDIDECRTIPACGDMNCV 60
OY 61 NONGVYLCPRTNPVYRGPSNPSTYSPGPPAAPPASNPPTISRLPCLICRFGYOMD 120
DB 61 NONGVYLCPRTNPVYRGPSNPSTYSPGPPAAPPASNPPTISRLPCLICRFGYOMD 120
OY 121 ESNOCVNDDECATDSHQCNPTOICITNTEGTYGCTDGYMLLEGQCLDIDECRYGCOOL 180
DB 121 ESNOCVNDDECATDSHQCNPTOICITNTEGTYGCTDGYMLLEGQCLDIDECRYGCOOL 180
OY 121 ESNOCVNDDECATDSHQCNPTOICITNTEGTYGCTDGYMLLEGQCLDIDECRYGCOOL 180
DB 121 ESNOCVNDDECATDSHQCNPTOICITNTEGTYGCTDGYMLLEGQCLDIDECRYGCOOL 180
OY 181 CANVGSVSCCTNPGFTLNEDGSCDVNMCATENPCVOCVNTYSGSFICGDCGYELE 240
DB 181 CANVGSVSCCTNPGFTLNEDGSCDVNMCATENPCVOCVNTYSGSFICGDCGYELE 240
OY 241 DGVHCSMDDECSEFELCOHECVNOGPTGYCSPPGYILLDDNRSODINECHRNHTCN 300
DB 241 DGVHCSMDDECSEFELCOHECVNOGPTGYCSPPGYILLDDNRSODINECHRNHTCN 300
OY 241 DGVHCSMDDECSEFELCOHECVNOGPTGYCSPPGYILLDDNRSODINECHRNHTCN 300
DB 241 DGVHCSMDDECSEFELCOHECVNOGPTGYCSPPGYILLDDNRSODINECHRNHTCN 300
OY 301 LQOTCYNLQGGFKCIDPIRCEPEYLRISDNRCMCPAENPCRCRDQPTILYRDMVSGRS 360
DB 301 LQOTCYNLQGGFKCIDPIRCEPEYLRISDNRCMCPAENPCRCRDQPTILYRDMVSGRS 360
OY 361 VPADIFQWQATRRYPGAYYIFQIKSGNEGREFYMRQGPISATLVMPRIKPREIQDL 420
DB 361 VPADIFQWQATRRYPGAYYIFQIKSGNEGREFYMRQGPISATLVMPRIKPREIQDL 420
OY 421 EMITVNTVNTVNTVNTVNTVNTVNTVNTVNTVNTVNTVNTVNTVNTVNTVNTVNT 448
DB 421 EMITVNTVNTVNTVNTVNTVNTVNTVNTVNTVNTVNTVNTVNTVNTVNTVNT 448

RESULT 3
FBL5_MOUSE STANDARD; PRT; 448 AA.
AC Q9WYH9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Ftbln-5 precursor (Ftbln-5) (Developmental arteries and neural crest
EGF-like protein) (Dance).
GN FBLN5 OR DANCE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

OX NCBI_TaxID=10090;
RN [1]
RP MEDLINE=99357779; PubMed=10428823;
RA Nakamura T., Ruiz-Lozano P., Lindner V., Yabe D., Taniwaki M.,
RA Furukawa Y., Kobuke K., Tashiro K., Lu Z., Andon N.L., Schaub R.,
RA Matsumori A., Sasayama S., Chien K.R., Honjo T.;
RT "DANCE, a novel secreted RGD protein expressed in developing,
RT atherosclerotic, and balloon-injured arteries.";
RL J. Biol. Chem. 274:22476-22483(1999).
CC -1- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH
CC INTERACTION OF INTEGRINS AND THE RGD MOTIF. COULD BE A VASCULAR
CC LIGAND FOR INTEGRIN RECEPTORS AND MAY PLAY A ROLE IN VASCULAR
CC DEVELOPMENT AND REMODELING.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF112151; AAD41767.1; -.
DR HSP: P00736; IAP0.
DR MGD: MGI:1346091; Fbln5.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR Pfam: PF00008; EGF_4.
DR SMART: SM00001; EGF_CA; 4.
DR SMART: SM00001; EGF_Like; 2.
DR PROSITE: PS00010; ASX_HYDROXYL; 4.
DR PROSITE: PS00022; EGF_1; FALSE_NEG.
DR PROSITE: PS01186; EGF_2; 4.
DR PROSITE: PS01187; EGF_CA; 6.
KW Repeat: EGF-like domain; Calcium-binding; Glycoprotein; Signal.
FT SIGNAL 1 23
FT CHAIN 24 448
FT DOMAIN 24 69 EGF-LIKE 1, DIVERGENT.
FT DOMAIN 127 167 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 168 206 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 207 245 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 247 287 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 288 333 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT SITE 54 56 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 131 144 BY SIMILARITY.
FT DISULFID 138 143 BY SIMILARITY.
FT DISULFID 155 166 BY SIMILARITY.
FT DISULFID 172 181 BY SIMILARITY.
FT DISULFID 177 190 BY SIMILARITY.
FT DISULFID 192 205 BY SIMILARITY.
FT DISULFID 211 221 BY SIMILARITY.
FT DISULFID 217 230 BY SIMILARITY.
FT DISULFID 222 245 BY SIMILARITY.
FT DISULFID 232 245 BY SIMILARITY.
FT DISULFID 251 262 BY SIMILARITY.
FT DISULFID 258 271 BY SIMILARITY.
FT DISULFID 273 286 BY SIMILARITY.
FT DISULFID 292 305 BY SIMILARITY.
FT DISULFID 299 314 BY SIMILARITY.
FT DISULFID 320 332 BY SIMILARITY.
FT CARBOHYD 283 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 448 AA; 50193 MM; F15CC70CCFBDC97 CRC64;

Query Match 94.9%; Score 2405; DB 1; Length 448;
Best Local Similarity 94.2%; Pred. No. 9.3e-176;
Matches 422; Conservative 10; Mismatches 16; Indels 0; Gaps 0;

OY 1 MFGIRLITVLTIALCLPSPGNNAOCTNFGDLDRSGOGLDIDECRTIPACGDMNCV 60

```

```

DR Genew; HGNC:3602; FBLSN.
DR MIM; 604580; -.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR Pfam: PF00008; EGF; 4.
DR SMART; SM00179; EGF_CA; 4.
DR PROSITE; PS00010; ASX_HYDROXYL; 4.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 4.
DR PROSITE; PS01187; EGF_CA; 6.
KW Repeat; EGF-like domain; Calcium-binding; glycoprotein; signal.
FT CHAIN 1 23 POTENTIAL.
FT SIGNAL 1 23 POTENTIAL.
FT DOMAIN 24 448 EGF-LIKE 1, DIVERGENT.
FT DOMAIN 127 167 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 168 206 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 207 246 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 247 287 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 288 333 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 333 356 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 54 56
FT DISULFID 131 144 BY SIMILARITY.
FT DISULFID 138 153 BY SIMILARITY.
FT DISULFID 155 166 BY SIMILARITY.
FT DISULFID 172 181 BY SIMILARITY.
FT DISULFID 177 190 BY SIMILARITY.
FT DISULFID 192 205 BY SIMILARITY.
FT DISULFID 211 221 BY SIMILARITY.
FT DISULFID 217 230 BY SIMILARITY.
FT DISULFID 232 245 BY SIMILARITY.
FT DISULFID 251 262 BY SIMILARITY.
FT DISULFID 273 286 BY SIMILARITY.
FT DISULFID 292 305 BY SIMILARITY.
FT DISULFID 299 314 BY SIMILARITY.
FT DISULFID 320 332 BY SIMILARITY.
FT CARBOHYD 283 296 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CONFLICT 69 70 IP -> HS (IN REF. 3).
FT CONFLICT 147 148 TE -> MS (IN REF. 3).
SQ SEQUENCE 448 AA; 50180 MW; 19FCA51FDA328003 CRC64;

```

```

Query Match 100.0%; Score 2533; DB 1; Length 448;
Best Local Similarity 100.0%; Pred. No. 1.8e-185;
Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MPQIKRLTYITIALCLPSPGNAOCTNGEDLDNOSGCLDIDECRTPEACRGDMCV 60
1 MPQIKRLTYITIALCLPSPGNAOCTNGEDLDNOSGCLDIDECRTPEACRGDMCV 60
QY 61 NONGYICIRTNVYRGPNYSPYSGPPAPAPPLSAPNYPTISAPLICRFGYOMD 120
61 NONGYICIRTNVYRGPNYSPYSGPPAPAPPLSAPNYPTISAPLICRFGYOMD 120
DB 61 NONGYICIRTNVYRGPNYSPYSGPPAPAPPLSAPNYPTISAPLICRFGYOMD 120
61 NONGYICIRTNVYRGPNYSPYSGPPAPAPPLSAPNYPTISAPLICRFGYOMD 120
QY 121 ESNQCVADVDCANDSHOCPQICINTEGTYCTDGYWLEGGCLDIDECRYGCOOL 180
121 ESNQCVADVDCANDSHOCPQICINTEGTYCTDGYWLEGGCLDIDECRYGCOOL 180
DB 121 ESNQCVADVDCANDSHOCPQICINTEGTYCTDGYWLEGGCLDIDECRYGCOOL 180
121 ESNQCVADVDCANDSHOCPQICINTEGTYCTDGYWLEGGCLDIDECRYGCOOL 180
QY 181 CANVPGSYSTCNGFTLNEDGRSCQDVNECATENPCVQTCVNTGSGFICRCDPGYLEE 240
181 CANVPGSYSTCNGFTLNEDGRSCQDVNECATENPCVQTCVNTGSGFICRCDPGYLEE 240
DB 181 CANVPGSYSTCNGFTLNEDGRSCQDVNECATENPCVQTCVNTGSGFICRCDPGYLEE 240
181 CANVPGSYSTCNGFTLNEDGRSCQDVNECATENPCVQTCVNTGSGFICRCDPGYLEE 240
QY 241 DGVHCSMDCECSFESEFLCOHECVNPGTYFCSPGTYILLDNRSQODINECHRHHTCN 300
241 DGVHCSMDCECSFESEFLCOHECVNPGTYFCSPGTYILLDNRSQODINECHRHHTCN 300
DB 241 DGVHCSMDCECSFESEFLCOHECVNPGTYFCSPGTYILLDNRSQODINECHRHHTCN 300
241 DGVHCSMDCECSFESEFLCOHECVNPGTYFCSPGTYILLDNRSQODINECHRHHTCN 300
QY 301 LQOTCYNLGGGFKCIDPRCEPEYLRISDNRCMPAENPGCRDOPFTILYRDMDVYSGRS 360
301 LQOTCYNLGGGFKCIDPRCEPEYLRISDNRCMPAENPGCRDOPFTILYRDMDVYSGRS 360
DB 301 LQOTCYNLGGGFKCIDPRCEPEYLRISDNRCMPAENPGCRDOPFTILYRDMDVYSGRS 360
301 LQOTCYNLGGGFKCIDPRCEPEYLRISDNRCMPAENPGCRDOPFTILYRDMDVYSGRS 360
QY 361 VPADIPQWQATRRPGAYYIRQISGNGREFFYMQTGPISTLVNTRPIKGPRTQLDL 420
361 VPADIPQWQATRRPGAYYIRQISGNGREFFYMQTGPISTLVNTRPIKGPRTQLDL 420

```

```

DB 361 VPADIPQWQATRRPGAYYIRQISGNGREFFYMQTGPISTLVNTRPIKGPRTQLDL 420
QY 421 EMITVNTVYINFRGSSVIRLRIYVSQYP 448
421 EMITVNTVYINFRGSSVIRLRIYVSQYP 448
DB 421 EMITVNTVYINFRGSSVIRLRIYVSQYP 448
421 EMITVNTVYINFRGSSVIRLRIYVSQYP 448

RESULT 2
FBLS_RAT STANDARD: PRT: 448 AA.
AC Q9WVH8; Q9R284;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fibulin-5 precursor (FBLN-5) (Developmental arteries and neural crest
DE EGF-like protein) (Dance) (Embryonic vascular EGF repeat-containing
DE protein) (EVEC).
GN FBLN5 OR DANCE.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99357779; PubMed=10428823;
RA Nakamura T., Ruiz-Lozano P., Lindner V., Yabe D., Tanikawa M.,
RA Furukawa Y., Kobuke K., Tashiro K., Lu Z., Andon N.L., Schaub R.,
RA Matsumori A., Sasayama S., Chien K.R., Honjo T.;
RT "DANCE, a novel secreted RGD protein expressed in developing,
RT atherosclerotic, and balloon-injured arteries.";
RL J. Biol. Chem. 274:22476-22483(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99278197; PubMed=10347091;
RA Kowal R.C., Richardson J.A., Milano J.M., Olson E.N.;
RT "EVEC, a novel epidermal growth factor-like repeat-containing protein
RT upregulated in embryonic and diseased adult vasculature.";
RL Circ. Res. 84:1166-1176(1999).
CC 1- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH
CC INTERACTION OF INTEGRINS AND THE RGD MOTIF. COULD BE A VASCULAR
CC LIGAND FOR INTEGRIN RECEPTORS AND MAY PLAY A ROLE IN VASCULAR
CC DEVELOPMENT AND REMODELING.
CC 1- SUBCELLULAR LOCATION: Secreted.
CC 1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DB EMBL; AF112153; AAD41769.1; -.
DR EMBL; AF137350; AAD25101.1; -.
DR HSSP; P00736; IAPQ.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR Pfam; PF00008; EGF; 4.
DR SMART; SM00179; EGF_CA; 4.
DR SMART; SM00001; EGF_LIKE; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 4.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 4.
DR PROSITE; PS01187; EGF_CA; 6.
KW Repeat; EGF-like domain; Calcium-binding; glycoprotein; signal.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 448 FIBULIN-5.
FT DOMAIN 24 69 EGF-LIKE 1, DIVERGENT.
FT DOMAIN 127 167 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 168 206 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 207 246 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 17:40:07 ; Search time 11.3157 seconds

(without alignments)
1642.086 Million cell updates/sec

Title: US-09-674-379A-13
Perfect score: 2533
Sequence: 1 MPGIKRLITVITLALCLPSP.....INFRGSSVIRLRIVSQYPF 448

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Swissprot_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2533	100.0	448	1 FBL5_HUMAN	Q9ubx5 homo sapien
2	2406	95.0	448	1 FBL5_RAT	Q9wv18 ratu
3	2405	94.9	448	1 FBL5_MOUSE	Q9wv18 mus musculu
4	1290	50.9	443	1 FBL4_CRIGR	O55058 cricetulus
5	1289	50.9	443	1 FBL4_HUMAN	O95967 homo sapien
6	1282	50.6	443	1 FBL4_MOUSE	O9wv19 mus musculu
7	1109.5	43.8	493	1 FBL3_RAT	O35568 ratu
8	1101	43.5	493	1 FBL3_HUMAN	O12805 homo sapien
9	778.5	30.7	684	1 FBL1_CHICK	Q13775 gallu
10	733.5	29.0	11221	1 FBL2_MOUSE	P37889 mus musculu
11	709.5	28.0	1184	1 FBL2_HUMAN	P98095 homo sapien
12	703.5	27.8	703	1 FBL1_MOUSE	O08879 mus musculu
13	698.5	27.5	703	1 FBL1_HUMAN	P23142 homo sapien
14	599.5	23.7	712	1 FBL1_CAEEL	O77469 caenorhadi
15	544	21.5	1394	1 LTRBS_HUMAN	P22064 homo sapien
16	544	21.5	1595	1 LTRBS_HUMAN	Q14766 homo sapien
17	533.5	21.0	2871	1 FBN1_MOUSE	P35555 homo sapien
18	531.5	21.0	2871	1 FBN1_BOVIN	P98133 bos tauris
19	525.5	20.7	2871	1 FBN1_PIG	O9t46 sus scrofa
20	523	20.6	1712	1 LTRB_RAT	O00918 ratu
21	518.5	20.5	2871	1 FBN1_MOUSE	O61554 ratu
22	518.5	20.5	2907	1 FBN2_MOUSE	O61555 homo sapien
23	517.5	20.4	2911	1 FBN2_HUMAN	P35555 mus musculu
24	475	18.8	956	1 MTN2_HUMAN	O00339 homo sapien
25	452	17.8	956	1 MTN2_MOUSE	O08766 mus musculu
26	421	16.6	931	1 EMRI_MOUSE	O61349 mus musculu
27	392	15.5	886	1 EMRI_HUMAN	Q14246 homo sapien
28	379	15.0	2470	1 NTC2_MOUSE	O35516 mus musculu
29	377	14.9	2471	1 NTC2_HUMAN	O04721 homo sapien
30	375	14.8	810	1 NTC1_HUMAN	O92832 homo sapien
31	372.5	14.7	816	1 NTC2_MOUSE	O61220 mus musculu
32	371.5	14.7	2471	1 NTC2_RAT	O9qw30 ratu
33	366	14.4	816	1 NTC2_HUMAN	O99435 homo sapien

34	364	14.4	810	1 NTC1_RAT	O62919 ratu
35	360.5	14.2	816	1 NTC1_CHICK	O90827 gallu
36	359.5	14.2	1964	1 NTC4_MOUSE	P31695 mus musculu
37	359	14.2	816	1 NTC2_RAT	O62918 ratu
38	356	14.1	2703	1 NTC1_DROME	P07207 drosophila
39	355.5	14.0	2437	1 NTC1_BRARE	P46530 brachydania
40	350.5	13.8	652	1 CD93_HUMAN	O9np93 homo sapien
41	347.5	13.7	2556	1 NTC1_HUMAN	P46531 mus sapien
42	347	13.7	644	1 CD93_MOUSE	O89103 mus musculu
43	346	13.7	1064	1 FBL1_STRPU	P10079 strongyloe
44	346	13.7	2524	1 NTC1_XENLA	P21783 xenopus lae
45	344	13.6	2321	1 NTC3_HUMAN	O9um47 homo sapien

ALIGNMENTS

RESULT 1
FBL5_HUMAN STANDARD; PRT; 448 AA.
ID FBL5_HUMAN
AC Q9UBX5; 075966;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fibulin-5 precursor (FBL-5) (developmental arteries and neural crest
EGF-like protein) (Dance) (Urine p50 protein) (UP50).
GN FBLN5 OR DANCE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RA Kostka G.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99357779; PubMed=10428823;
RA Nakamura T., Ruiz-Lozano P., Lindner V., Yabe D., Taniwaki M.,
RA Furukawa Y., Kobuke K., Tashiro K., Lu Z., Andon N.L., Schaub R.,
RA Matsumori A., Sasayama S., Chien K.R., Honjo T.;
RT "DANCE", a novel secreted RGD protein expressed in developing,
atherosclerotic, and balloon-injured arteries.";
RL J. Biol. Chem. 274:22476-22483(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Urine;
RA Zemel R., Sholto O., Shaul Y.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH
INTERACTION OF INTEGRINS AND THE RGD MOTIF. COULD BE A VASCULAR
LIGAND FOR INTEGRIN RECEPTORS AND MAY PLAY A ROLE IN VASCULAR
DEVELOPMENT AND REMODELING.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN HEART, OVARY, AND
COLON BUT ALSO IN KIDNEY, PANCREAS, TESTIS, LUNG, AND PLACENTA.
CC NOT DETECTABLE IN BRAIN, LIVER, THYMUS, PROSTATE, OR PERIPHERAL
BLOOD LEUKOCYTES.
CC -!- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use, by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ133490; CAB38568.1; -;
DR EMBL: AF112152; AAD41768.1; -;
DR EMBL: AF093118; AAC62107.1; -;
DR HSSP: P00736; IAPQ.

```
OY 217 -----CVO-----TCVNTYGSFICRCDPGYELEDGVHC 245
Db 1392 YRCLCYDGFMASEDMKTCVDVNECDLNPNICLSGTCENTKGSFICHCDMGSGKRGKTCG 1451
OY 246 SDMECSSEFELC-QHE-CVNOPGYFCSPGPIILLDNRSCODINECEHRNHTCNLQ 303
Db 1452 TDINCEIGAHNCGRAVCTNAGSFKCSGPGMI--GDGICKTDLDECSNGTMCQSHA 1509
OY 304 TCYNLQGGFKCIDPIRCEPYL-----RISDRCMCPAENPGCRDOP 345
Db 1510 DCKNTMGSYRCL-----CKEGYTGDFCTDLDCESENILNLC--GNGOCLNAP 1555
```

RESULT 15

A55567

fibrillin 1 - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 02-Aug-2002

C:Accession: A55567

R:Listra, D.J.; Li, L.; Potter, K.A.; Womack, J.; Byers, P.H.

Genomics 23, 480-485, 1994

Title: Sequence of the coding region of the bovine fibrillin cDNA and localization to

Reference number: A55567; MUID:95137597; PMID:7835900

Accession: A55567

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-2871 <Till>

C:Cross-references: GB:L28748; NID:9508427; PIDN:AAA74122.1; PID:9508428

C:Superfamily: fibrillin 1; EGF homology

F:1201-1236/Domain: EGF homology <EGF>

Query Match 21.0%; Score 531.5; DB 2; Length 2871;

Best Local Similarity 29.6%; Pred. No. 4,4e-27;

Matches 122; Conservative 48; Mismatches 129; Indels 113; Gaps 16;

```
OY 16 CLPSPGNAOAGCTNGFDLDQSGCLDIDECRTIPACRGDMCVQNGGYLCIPRTNPV 75
Db 1044 CRNTIGSFRCRCDGFDLSEERNCTDIDECRISPLDC--GRGQCVNTPGDEFC--KCDEG 1100
OY 76 YRGPSNPSYSTRSGPYPAAPPLSAPNPTISR---PLICRGTQMDSE----- 122
Db 1101 YESGF-----MMKNCMDIDECQRPDLRCRGVCIINTEGYSYRCGCP 1142
OY 123 -----NOCVDVDECATDSHCNPTQICINTEGGYCSCTDGYWLEGO--CLDIDEC 172
Db 1143 GHQLAPNLSACIDINECELSAHLG--PHGRCVNLIGKYQCAQNGYHSTPRLPCVDIDEC 1201
OY 173 RY--GYCOQLCANVPQSYSCTCNPGFTLNEDGRSCODVNECATENP----- 216
Db 1202 SIMNGCETFTNBSGYSYCSQPGFALMPDQRSCTDIDEC--EDNPNICDGGQCTNIPGE 1260
OY 217 -----CVO-----TCVNTYGSFICRCDPGYELEDGVHC 245
Db 1261 YRCLCYDGFMASEDMKTCVDVNECDLNPNICLSGTCENTKGSFICHCDMGSGKRGKTCG 1320
OY 246 SDMECSSEFELCQHE--CVNPGYFCSPGPIILLDNRSCODINECEHRNHTCNLQ 303
Db 1321 TDINCEIGAHNCGRAVCTNAGSFKCSGPGMI--GDGICKTDLDECSNGTMCQSHA 1378
OY 304 TCYNLQGGFKCIDPIRCEPYL-----RISDRCMCPAENPGCRDOP 345
Db 1379 DCKNTMGSYRCL-----CKEGYTGDFCTDLDCESENILNLC--GNGOCLNAP 1424
```

Search completed: July 3, 2003, 18:22:41
Job time : 20.0023 secs

RESULT 11

T43210

fibulin-1D precursor - Caenorhabditis elegans (fragment)

C:Species: Caenorhabditis elegans

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 02-Sep-2000

C:Accession: T43210

R:Barth, J.L.; Argaves, K.M.; Roark, E.F.; Little, C.D.; Argaves, W.S.

submitted to the EMBL Data Library, June 1998

A:Description: Identification of chicken and C. elegans fibulin-1 homologs and character

A:Reference number: Z2337

A:Accession: T43210

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-589 <BAR>

A:Cross-references: EMBL:AF070477; PIDN:AAC24035.1

C:Genetics:

A:Note: Intron positions not resolved (incomplete sequence)

C:Superfamily: unassigned EGF-related proteins; EGF homology

Query Match 23.5%; Score 596; DB 2; Length 589;

Best Local Similarity 27.3%; Pred. No. 6,3e-32;

Matches 159; Conservative 53; Mismatches 163; Indels 208; Gaps 21;

15 ICLPSPG-NAOACTNGFDLDROSGOCLD-IDECRTIPACRGDMCMVONGGYLCIPRT 72

66 LCHDRGGEKVECSRSRSGFDLAPDGMACVDRHIDECATLMDCLLESQRCLNTPGSFKCI 122

73 NPVYRGPYNPSTYPSGPRPAAPPLASAPNPTISRLICRFGYQMD-ESNOCVDVDEC 131

123 -----RTLSGCTGTAADSETERNNCFLLIINTFNCKYFEV 146

132 ATDSHOCNPTQICINTEGGYTC-----SCTDGYWLEGGOCLD 170

147 NLGSHCCGGLYOCRNTOGSRCDARKKGDELQNPMTGECSTICPNYIKNGMCNID 206

171 ECRYGT-----CQO---LCA----- 182

207 ECVTGHCNCGAGECVNTPGSFRCQKGNLCAHGEVNGATGFCEDVNECQGVGSMECT 266

183 NYPGSGSCNPGFTLNE----- 200

267 NLPGRKCKGPGYEENDAKKRCEDEDECIFAGHVCDDLSEACINTIGSFCKCKPGFOL 326

201 --DGRSCQDVNECATE-NPCVOTCVNTYGSFICRCDPGYLEEDGVHCSMDDECSF 253

327 ASDGRCEVNECTTGIAACEQKCVNIPGSYQCICDGFALGPGSTKEDIDECSTMG 386

254 SEFLCOHECVNPGTYFSCSPGTYLLDNRSCODINECEHNRHTCNLQOTCYNLQGGFK 313

387 GMDLCMGCCINTKGSYLCCPGRKIQPDGRTCVDVDECA-MGECAGSDKVCVNTIGSFK 445

314 CIDPIRCEPRLISNR-----C-MCPAENPGC-RDQPTILYRMDVYSGRSV 361

446 C-HSIDCPNTYINDSLNKNOIADGYSCTIKVCSPTEDTCLGNHTREVLIOFRAVPSLKTII 504

362 -PADI-----FOMQATRYPGAYYIFOLKSGNEGRFYMROQTPISATLVMPRI 410

505 SPIEVSRIYTHMGVPSVDNLDYVQGRHRIYQERNIG-----IYQLVAPI 551

411 KGPRIQLDLEMTVNTVINEGSSVIR-----LRIYSQYRF 448

552 SGP-----YVERIKVINHTKSRGTGVLAFNEAIEISVSKYRF 589

RESULT 12

T22793

hypothetical protein F56H11.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Aug-2002

C:Accession: T22793; T24489

R:Wilkinson, J.

submitted to the EMBL Data Library, January 1996

A:Reference number: Z19616

A:Accession: T22793

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-798 <WIL>

A:Cross-references: EMBL:Z68749; PIDN:CAA92962.1; GSPDB:GN00022; CESP:F56H11.1

R:Lloyd, C.

submitted to the EMBL Data Library, December 1995

A:Reference number: Z19697

A:Accession: T24489

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-798 <WIL>

A:Cross-references: EMBL:Z68219; PIDN:CAA92483.1; GSPDB:GN00022; CESP:F56H11.1

A:Experimental source: clone T05A1

C:Genetics:

A:Gene: CESP:F56H11.1

A:Map position: 4

A:Introns: 14/1; 92/3; 144/1; 195/1; 281/1; 325/2; 371/2; 390/1; 437/1; 471/3; 498/3;

C:Superfamily: fibulin-1; EGF homology

Query Match 22.8%; Score 577.5; DB 2; Length 798;

Best Local Similarity 26.7%; Pred. No. 1.4e-30;

Matches 162; Conservative 50; Mismatches 180; Indels 215; Gaps 23;

15 ICLPSPG-NAOACTNGFDLDROSGOCLDIDECRTIPACRGDMCMVONGGYLC----- 68

167 LCHDRGGEKVECSRSRSGFDLAPDGMACVDRHIDECATLMDCLLESQRCLNTPGSFKCI 226

69 -IP--RTNPVYRGP--YSNPSTYPSGPRPAAP----- 98

227 RLVPYHRRANRIGNAPRRMRDPPYSR--AGEYREASQANTERGCWGLFQGHCVDIDE 284

99 -----LSAPNPTISRLICRFGYQMD-----ESNOCV----- 126

285 CATLMDCLLESQRCLNTPGSFKCIPTLSGCTGTADSETERNNCFLLIINTFNCKYFEV 344

127 -DVDECATDSHOCNPTQICINTEGGYTC-----SCTDGYWLEGGOCLD 154

345 EDVDECMGSHDGPPLVOCRTQGSYRCDARKKGDELQNPMTGECSTICPNYIKNGMCNID 404

155 -----CTDGYWL--LEGGCLDIDECRYGCOOL-CANVPGSYCTCMP 194

405 ECYNTPGSFRCQKGNLCAHGEVNGATGFCEDVNECQGVGSMECTNLPETYKCKGCP 464

195 GFTLNE----- 210

465 GYEFNDAKKRCEDVDECIFAGHVCDDLSEACINTIGSFCKCKPGFOLASDGRCEVNE 524

211 CATE-NPCVOTCVNTYGSFICRCDPGYLEEDGVHCSMDDECSF-----SEFLCOHECVNQ 265

525 CTGIAACEQKCVNIPGSYQCICDGFALGPGSTKEDIDECSTMG 584

266 PGTYFCSGPGTYLLDNRSCODINECEHNRHTCNLQOTCYNLQGGFKCIDPIRCEPYL 325

585 KGSYLCQCPGPKYKIQPDGRTCVDVDECA-MGECAGSDKVCVNTIGSFKC-HSIDCPNTYI 642

326 RISDNR-----CMCPAENPGCRDQPTILYRMDVYSGRSVP----- 362

643 HDSLNNKRCNRPSACGILPEE--CSKVPLELTJTOFISL--ARAVPISSHRAPIYLFKVS 697

363 ----ADI-----FOMQATRYPGAYYIFOLKSGNEGRFYMROQTP--ISATLVMPRIKGP 413

698 APNHADLEVNFELQKTTIYGAPVLAIRAN-----FLQNGEKRNASVYVILRSLDGP 752

RESULT 13

A35626

hypothetical protein F56H11.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Aug-2002

C:Accession: T22793; T24489

R:Wilkinson, J.

submitted to the EMBL Data Library, January 1996

A:Title: fibulin, a novel protein that interacts with the fibronectin receptor beta-subunit
 A:Reference number: A32826; MUID:89354537; PMID:2527614
 A:Accession: A32826
 A:Molecule type: protein
 A:Residues: 30-35,'SX',38-40,'SH',43-44 <AR3>
 C:Genetics:
 A:Gene: GDB:FBLN1; FBLN
 A:Cross-references: GDB:278285; OMIM:135820
 A:Map position: 22q13.3-22q13.3
 C:Superfamily: fibulin-1; EGF homology
 C:Keywords: alternative splicing; glycoprotein
 F:1-29/Domain: signal sequence #status predicted <Sig>
 F:30-683/Product: fibulin 1 splice form C #status predicted <MAT>
 F:180-214/Domain: EGF homology <EGF>
 F:485-523/Domain: EGF homology <EGF1>
 F:59,535/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 29.1%; Score 737; DB 2; Length 683;
 Best Local Similarity 35.4%; Pred. No. 4,1e-41;
 Matches 157; Conservative 67; Mismatches 154; Indels 66; Gaps 17;

```

QY 16 CLSPGNAQACTNGFDLRSGGCLDIDECRTPEACRGDMCVNNGYLCTPRTNPV 75
DB 288 CRP-----KLCQSGFLOD-ALGNCIDINECLISAPCPICHTCINTGSGSYTC----- 334
QY 76 YRGPYSNPYSTPYSGPYPAAPPLSAPNYPTISRLPGRFYGM-DESNOCVDVDECATD 134
DB 335 -----QKNVFN-----CGRGTHLNEBGRVDEVCAP 363
QY 135 SHOCNPTQICINTEGGYTCSTDGWYL-LEGOCLDIDEC-RY-GYCOOLCANVGSYS 189
DB 364 AEECGKGRHCVNSGSPRCCKTGTFDGIISRMCVDVNECQRYRPLRGHNCENTLGSTL 423
QY 190 CTGNPGFTLNEBDRSCODVNECAENPCVQTCVNTYGSFTRCDPGLYLEE-DGVHCSDM 248
DB 424 CSOSVGRSLVDGSCDINECSS-SPSCQECANVGSYQCYCRGQYLDSDVDTCEDI 482
QY 249 DECSF--SEELCOHECVNOPTGYFCSCP-PGYTLIDNRSQDINECHRNHTCNLOOTC 305
DB 483 DECALPTGHCIGYRCINIGSFQSCSPSSGYRLAPNRKQDIDECVTGTHNCINSEC 542
QY 306 YNIGGFKCIDPTICEEPYRLISDNRC-MCPA-ENPGCRDPFTILYRDMVYSGRSVPA 363
DB 543 FNIQGAERCL-AEECPENYRRSATRCERLPCHENRRECSKPLRLTYVHLSFPTIQAPA 601
QY 364 DIFOMATRRYPGAYYIFQIKSGNEGREFTRMGCPISATLVMTIRPKGPREIQLDLEMI 423
DB 602 VYFRMGPSSAVPEDSQMLATGGNEGEFTTRKVPSPSGVALTKPVEPRDL-----LL 656
QY 424 TVNTVINFNG--SSVIRLRIVYS 444
DB 657 TVKMDLSRHGTVSSFAKLFIFVS 680

```

RESULT 5
 A:Accession: A49457
 A:Molecule type: mRNA
 A:Residues: 1-1221 <PAN>
 A:Cross-references: GB:X75285; NID:9437046; PIDN:CAAS3040.1; PID:9437047
 R:Sasaki, T.; Mann, K.; Murphy, G.; Chu, M.L.; Timpl, R.
 Eur. J. Biochem. 240, 427-434, 1996
 A:Title: Different susceptibilities of fibulin-1 and fibulin-2 to cleavage by matrix metalloproteinases
 A:Reference number: A49457; MUID:94064787; PMID:8245130
 A:Accession: A49457
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1221 <PAN>
 A:Cross-references: GB:X75285; NID:9437046; PIDN:CAAS3040.1; PID:9437047
 R:Sasaki, T.; Mann, K.; Murphy, G.; Chu, M.L.; Timpl, R.
 Eur. J. Biochem. 240, 427-434, 1996
 A:Title: Different susceptibilities of fibulin-1 and fibulin-2 to cleavage by matrix metalloproteinases
 A:Reference number: S74094; MUID:96439073; PMID:8841408

A:Accession: S74095
 A:Molecule type: protein
 A:Residues: 236-238,'X',240-247;260-275;336-344,'L',346-361;405-426;566-568,'EW',569-570
 C:Superfamily: fibulin-2; EGF homology
 C:Keywords: calcium binding; duplication; extracellular matrix; glycoprotein; homotrimer
 F:942-978/Domain: EGF homology <EGF>

Query Match 29.0%; Score 733.5; DB 2; Length 1221;
 Best Local Similarity 35.3%; Pred. No. 1.2e-40;
 Matches 146; Conservative 60; Mismatches 149; Indels 59; Gaps 11;

```

QY 23 AQACQTNAGFDLRSGGCLDIDECRTPEACRGDMCVNNGYLCTPRTNPVYRGSYN 82
DB 829 ARQRCMDGF-LQDEGNCVDINECTSLIEPERSGFSINTGSGSYTC----- 873
QY 83 PYSTPYSGPYPAAPPLSAPNYPTISRLPGRFYGM-DESNOCVDVDECATDHSQNP 141
DB 874 -----QKNVFN-----CGRGTHLNEBGRVDEVCAP 363
QY 142 QICINTEGGYTCSTDGWYL-LEGOCLDIDECRYG--YCOOLCANVGSYSCTCNPGF 196
DB 910 QLCYNLPGSYRCDCKPGFQDAFGRTCIDVNECAVSPGLCQHNCENTPGSYRSCAAGF 969
QY 197 TLNEBDRSCODVNECAENPCVQTCVNTYGSFTRCDPGLYLEE-DGVHCSDMDECS-FSE 255
DB 970 LLAADGHCEDVNECETRR-CSQECANVGSYQCYCRGQYLAADGHTCTDIDECQAG 1028
QY 256 FLCQHCNVNOPTGYFCSCP-PGYTLIDNRSQDINECHRNHTCNLOOTCYNLOGGFKC 314
DB 1029 ILCTFFCVNPGSYQACCPGQGTMTMANGSKRLDECALGTHNCSEMETHTNIOGSRIC 1088
QY 315 IDPRCEEPYRLISDNRCMCPAENPGCRD-----QPTILYRDMVYSGRSVPA 368
DB 1089 L-RFDCPPNVYRVSQKRC---ERTTCQDITECQSPARLHYQLNFTQGLVLAHPIRI 1143
QY 369 QATTRYPGAYYIFQIKSGNEGREFTRMGCPISATLVMTIRPKGPREIQLDLEMI 422
DB 1144 GRAPAFAGDTISLITTKGNEGEFTTRKVPSPSGVALTKPVEPRDL-----LL 1197

```

RESULT 6
 A:Accession: A55184
 A:Molecule type: protein
 N:Alternate names: protein DKFZp586A1519.1
 C:Species: Homo sapiens (man)
 C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 02-Aug-2002
 C:Accession: A55184; T08744
 R:Zhang, R.Z.; Pan, T.C.; Zhang, Z.Y.; Mattei, M.G.; Timpl, R.; Chu, M.L.
 Genomics 22, 425-430, 1994
 A:Title: Fibulin-2 (FBLN2): human cDNA sequence, mRNA expression, and mapping of the gene
 A:Reference number: A55184; MUID:95104855; PMID:7806230
 A:Accession: A55184
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1184 <ZHA>
 A:Cross-references: GB:X82494; NID:9575232; PIDN:CA57876.1; PID:9575233
 R:Wandut, R.; Heudner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, March 1999
 A:Reference number: Z16471
 A:Accession: T08744
 A:Molecule type: mRNA
 A:Residues: 656-719,'QDECLMGADCSRQPCVNTLGSFYCVNHTVLCADGTYILNAHKCYD',720-853,'T',855-856
 A:Cross-references: EMBL:AD050095
 A:Experimental source: adult uterus; clone DKFZp586A1519
 C:Genetics:
 A:Gene: GDB:FBLN2
 A:Cross-references: GDB:293037; OMIM:135821
 A:Map position: 3p25-3p24
 A:Note: DKFZp586A1519.1
 C:Superfamily: fibulin-2; EGF homology
 C:Keywords: alternative splicing; extracellular matrix
 F:1-27/Domain: signal sequence #status predicted <Sig>
 F:28-1184/Product: fibulin-2 protein #status predicted <MAT>

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 17:36:12 ; Search time 18.0023 Seconds
(without alignments)
2392.374 Million cell updates/sec

Title: US-09-674-379A-13

Perfect score: 2533

Sequence: 1 MPQIKRLITVITIALCLPSP.....INFRGSYRLRLRYVSQYPF 448

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR_73:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1109.5	43.8	493	2 JC5621	epidermal growth f
2	963.5	38.0	387	2 I38449	extracellular prot
3	746	29.5	685	2 S78040	fibulin, splice fo
4	737	29.1	683	2 C36346	fibulin-1 precursor
5	733.5	29.0	1221	2 A49457	fibulin-2 precursor
6	709.5	28.0	1184	2 A5184	fibulin-2 precursor
7	703.5	27.8	705	2 S34968	fibulin, splice fo
8	619	24.4	601	2 B36346	fibulin-1 precursor
9	606.5	23.9	689	2 T42760	fibulin, splice fo
10	599.5	23.7	712	2 T42990	fibulin-1, splice
11	596	23.5	589	2 T43210	fibulin-1D precurs
12	577.5	22.8	798	2 T22793	hypothetical prote
13	544	21.5	1394	2 A35626	transforming grow
14	532.5	21.0	3002	2 A47221	fibritillin-1 precu
15	531.5	21.0	2871	2 A55567	fibritillin-1 - bov
16	523	20.6	1712	2 A58261	masking protein pr
17	518.5	20.5	2871	2 A56624	fibritillin-1 precu
18	518.5	20.5	2907	2 A57278	fibritillin-2 precu
19	517.5	20.4	2918	2 A54105	latent transformin
20	480.5	19.0	1820	2 A55494	latent transformin
21	475	18.8	741	2 T46488	hypothetical prote
22	461	18.2	1251	2 A57293	latent transformin
23	459	18.1	1620	2 T27283	hypothetical prote
24	444	17.5	1574	2 T13954	MEGF6 protein - ra
25	392	15.5	886	2 A57172	probable hormone r
26	388	15.3	3507	2 T34513	hypothetical prote
27	372.5	14.7	1106	2 T18739	hypothetical prote
28	371.5	14.7	2471	2 A49128	cell-fate determin
29	365.5	14.4	1203	2 A49175	Notch-1 protein -

30	364	14.4	810	2 T10756	Nel-homolog protei
31	361	14.3	1081	2 T31329	receptor tyrosine
32	359.5	14.2	1964	2 T09059	notch4 - mouse
33	359	14.2	511	2 T17298	hypothetical prote
34	358	14.1	2555	2 A40043	notch protein homo
35	356	14.1	2703	1 A24420	notch protein - fr
36	355.5	14.0	2437	2 A24612	transmembrane prot
37	346	13.7	1064	2 A40136	fibropellin Ia - s
38	346	13.7	2524	2 A35844	notch protein - Af
39	344	13.6	2321	2 S78549	notch3 protein - h
40	340.5	13.4	673	2 A48089	growth arrest-spec
41	337.5	13.3	2318	2 S45306	notch 3 protein -
42	337.5	13.3	2531	2 S18188	notch protein homo
43	337	13.2	2531	2 T31070	notch homolog - se
44	335.5	13.2	674	2 I54476	growth potentialin
45	335	13.2	2531	2 A46019	Notch-1 protein -

ALIGNMENTS

RESULT 1

JC5621
epidermal growth factor-like protein, T16 precursor - rat

C:Species: Rattus norvegicus (Norway rat)
C:Date: 09-Oct-1997 #sequence_revision 07-Nov-1997 #text_change 05-Nov-1999

C:Accession: JC5621
R:Oraki, T.; Kondo, K.; Nakamura, Y.; Ichimura, S.; Nakagawara, A.; Sakiyama, S.

Biochem. Biophys. Res. Commun. 237, 245-250, 1997
A:Title: Interaction of DA41, a DAN-binding protein, with the epidermal growth factor.

A:Reference number: JC5621; MUID:97415782; PMID:9268694
A:Accession: JC5621

A:Molecule type: mRNA
A:Residues: 1-493 <OZA>

A:Cross-References: DDBJ:D89730; NID:92429082; PIDN:BA22265.1; PID:dt023127; PID:92

C:Comment: This protein plays a role in the regulation of cell growth by interacting

C:Keywords: glycoprotein
F:1-17/Domain: signal sequence #status predicted <Sig>

F:28-70,158-199,200-237,238-277,278-318,319-359/Region: epidermal growth factor-like

F:249/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 43.8%; Score 1109.5; DB 2; Length 493;
Best local similarity 42.3%; Pred. No. 1.1e-65;

Matches 206; Conservative 73; Mismatches 161; Indels 47; Gaps 5;

QY	7	ILTYTIALCLPSPENAGACCTNGRDLDRQSQCDDIDECRTPEACRGMVQNGCY	66
DB	9	MTLALVKSQVTEETITTYTCTDGYEMDPVROCKDIDCDIAPDCKGMRKCVNHGYG	68
QY	67	LCIPRTNVPYRGPSNPYSTPS-----GPPAAPPASAP	102
DB	69	LCLEPKTAQIIYNNEDPQOETPAEASSGAATGTIAARSMATSGVITGGGFIATAVAGP	128
QY	103	NYPT-----ISRLPCHFGYQMDSESNOCVDVDECATDSHCNPQT	142
DB	129	EVQGRNNFVIRBNPADPQRIPSNPSSHRQCAAGYQSHNVCQDIDECSTGHNRCLDQ	188
QY	143	ICINTEGCTSCIDGYMLBEGCQDIDECRI-GYCOQLCANVPGRYSCTCNGFTLNED	201
DB	189	VCINLRGSFTCHCLPQYKRGECVADIDECSSVPPYCHQCQVTPGSGFYCQCNFGFOLAN	248
QY	202	GRSCQDVACATENPCVQTVNTVGSFTCRCPGYELBEGVHCSMDSESFEEQHE	261
DB	249	NYTCVDINECDASNCQACQCVITLDSFTCCQNGVYELSDRLNCEDIDECRTSSYLQCYQ	308
QY	262	CVNPGTFYSCPPYILLDNRSQCQDINECHRNHTCMLOQTCVMLOGGFKCIDIRCE	321
DB	309	CVNPGKFSKCMCPQGYQYV-RSRTQDINECTTNE-CAEDECMNHYHGFRRYPQPCQ	366
QY	322	EPYLRISDNRCMCAENPGCRDQPTIYRDMVVGSRSPADIFQMAATRYGAYYIF	361
DB	367	DPYVLTSENRCVCPVSNMCRDVPQSIYKYVNNISDRSVPSDIFQIAQTIIYANTINIF	426

XX (ONCY) ONO PHARM CO LTD.
XX

PI Honjo T, Tashiro K, Nakamura T;
XX

DR WPI: 2000-038646/03.
XX

DR N-PSDB: AAZ39386, AAZ39387.
XX

PT Novel mouse polypeptides for treatment, e.g. arteriosclerosis and myoma
XX

PS Claim 1; Page 54-56; 70pp: Japanese.
XX

CC The invention provides mouse polypeptides for treatment of diseases due
CC to abnormal proliferation of smooth muscle. The polypeptides can be
CC produced by standard recombinant methodology. The polypeptides can be
CC used according to their inhibition of the proliferation of vascular
CC smooth muscle cells, particularly in treating arteriosclerosis or re-
CC narrowing by vascular endothelial thickening after percutaneous
CC transluminal coronary angioplasty (PTCA), or myoma, hematopoietic cell-
CC regulatory activity, cytokine activity, tissue generation/repairation
CC activity, actin/inhibitin activity, taxis and chemotaxis activity, blood
CC coagulation/thrombotic activity, receptor/ligand activity, cadonelin/
CC tumour metastasis inhibiting activity; tumor inhibition, and as nutrient.
CC The present sequence represents the protein of the invention which can be
CC used for modulating smooth muscle cell proliferation.
XX

SQ Sequence 461 AA:

Query Match 95.6%; Score 2302; DB 21; Length 461;
Best Local Similarity 94.8%; Pred. No. 2.5e-153;

Matches 401; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

OY 1 OCTNGFDLDROSGGQCLDIDECRTIPACRGDMCMVONGGYLCIPRTNPYRGPSNPS 60
DB |||||||
39 OCTNGFDLDROSGGQCLDIDECRTIPACRGDMCMVONGGYLCIPRTNPYRGPSNPS 98
OY 61 TPYSGPYPAAPPLSANPTTISRPLICRGYOMDESNOCVDECATDSHCNPTQICI 120
DB |||||||
99 TSYSGPYPAAPPLSANPTTISRPLICRGYOMDESNOCVDECATDSHCNPTQICI 158
OY 121 NTEGGYTCSTGCVWLLLEGQCLDIDECRYGYCOQLCANVPGSYSCNPGFTLNEDGRSC 180
DB |||||||
159 NTEGGYTCSTGCVWLLLEGQCLDIDECRYGYCOQLCANVPGSYSCNPGFTLNEDGRSC 218
OY 181 ODVNECATPENCVCVNTYSGFICRCDDRGYLEEDGVHCSMDDECSFSEFLCOHECVNQ 240
DB |||||||
219 ODVNECATPENCVCVNTYSGFICRCDDRGYLEEDGVHCSMDDECSFSEFLCOHECVNQ 278
OY 241 PGTYFCSCPFGYLLDDNRSQDINCEHNRHTCNLOQTCYNLOGGFKCIDPISCEEPYL 300
DB |||||||
279 PGTYFCSCPFGYLLDDNRSQDINCEHNRHTCNLOQTCYNLOGGFKCIDPISCEEPYL 338
OY 301 RISDNRCMCPAENPCGRDQPTLLYRDMDVSGRSVPADIFOMQATTRYPGAYYIFQIKS 360
DB |||||||
339 LIGENRCMCPAENPCGRDQPTLLYRDMDVSGRSVPADIFOMQATTRYPGAYYIFQIKS 398
OY 361 GNEGREFYMRQTSATLYMTRPIKGPRIQDLEMITVNTVYNFSGSSYIRLRIYYSQ 420
DB |||||||
399 GNEGREFYMRQTSATLYMTRPIKGPRIQDLEMITVNTVYNFSGSSYIRLRIYYSQ 458
OY 421 YPF 423
DB |||
459 YPF 461

Search completed: July 3, 2003, 17:54:48
Job time : 30.5959 secs

DT 22-FEB-2000 (first entry)
 XX Smooth muscle proliferation modulating protein mature sequence.
 XX
 XX Cell proliferation; vascular; smooth muscle cell; arteriosclerosis; PTCA;
 XX endothelial thickening; percutaneous transluminal coronary angioplasty;
 XX myoma; hematopoietic cell-regulation; cytokine; tissue generation; taxis;
 XX actin activity; chemotaxis; blood coagulation; thrombotic; tumor; mouse;
 XX metastasis; nutrient.
 XX
 OS Mus musculus.
 XX
 XX WO9955863-A1.
 XX
 XX 04-NOV-1999.
 XX
 XX 28-APR-1999; 99WO-JP02283.
 XX
 XX 28-APR-1998; 98JP-0119731.
 XX
 XX (ONOX) ONO PHARM CO LTD.
 XX
 XX Honjo T, Tashiro K, Nakamura T;
 XX WPI: 2000-038646/03.
 XX N-PSDB; AA239388.
 XX
 XX Novel mouse polypeptides for treatment, e.g. arteriosclerosis and myoma
 XX
 XX Claim 1; Page 63-65; 70pp; Japanese.
 XX
 CC The invention provides mouse polypeptides for treatment of diseases due
 CC to abnormal proliferation of smooth muscle. The polypeptides can be
 CC produced by standard recombinant methodology. The polypeptides can be
 CC used according to their inhibition of the proliferation of vascular
 CC smooth muscle cells, particularly in treating arteriosclerosis or re-
 CC narrowing by vascular endothelial thickening after percutaneous
 CC transluminal coronary angioplasty (PTCA), or myoma, hematopoietic cell-
 CC regulatory activity, cytokine activity, tissue generation/repatriation
 CC activity, actin/inhibitin activity, taxis and chemotaxis activity, blood
 CC coagulation/thrombotic activity, receptor/ligand activity, cadohelin/
 CC tumour metastasis inhibiting activity; tumor inhibition, and as nutrient.
 CC The present sequence represents the mature protein of the invention which
 CC can be used for modulating smooth muscle cell proliferation.
 CC
 XX Sequence 423 AA:
 SO
 Query Match 95.6%; Score 2302; DB 21; Length 423;
 Best Local Similarity 94.8%; Pred. No. 2.2e-153;
 Matches 401; Conservative 9; Mismatches 13; Indels 0; Gaps 0;
 QY 1 OCTNGFDLDRSGQCIDIDECPRIPEACRGDMACVNONGGYLCIPRTNRYRGPSNPS 60
 DB 1 OCTNGFDLDRSGQCIDIDECPRIPEACRGDMACVNONGGYLCIPRTNRYRGPSNPS 60
 QY 61 TPYSGPYAAAPLSPAPNTISRPLICRGYOMDESNOCVDECATDSHOCPTQICI 120
 DB 61 TSYSPFPAAPAPVPRASNTPTISRPLVCRGMYOMDEBNOCVDECATDSHOCPTQICI 120
 QY 121 NTEGGYTCSTGTYWLEGGQCIDIDECPRIPEACRGDMACVNONGGYLCIPRTNRYRGPSNPS 180
 DB 121 NTEGGYTCSTGTYWLEGGQCIDIDECPRIPEACRGDMACVNONGGYLCIPRTNRYRGPSNPS 180
 QY 181 ODVNECATENPCVQTCVNTYSGFTICRCDPEYELBEDGVCSDMDECSFSEFLCOHECVNQ 240
 DB 181 ODVNECATENPCVQTCVNTYSGFTICRCDPEYELBEDGVCSDMDECSFSEFLCOHECVNQ 240
 QY 241 PGTYSCPPGYIILDDNRSQDINCEHNRHNCNLOQTQYNLOGGKCIDIPICEEPYL 300
 DB 241 PGTYSCPPGYIILDDNRSQDINCEHNRHNCNLOQTQYNLOGGKCIDIPICEEPYL 300
 QY 301 RISDNRCMCPAENPCRDQPFITLYRDMVVSGRSVADIFQOMATRRYGAAYYIFQIKS 360

DB 301 LIGENRCMCPAENPCRDQPFITLYRDMVVSGRSVADIFQOMATRRYGAAYYIFQIKS 360
 QY 361 GNEGREFYMRQGPISATLVMTRPIKGPRIQDLEMTVNTVNFGRSSVIRLRIVYQ 420
 DB 361 GNEGREFYMRQGPISATLVMTRPIKGPRIQDLEMTVNTVNFGRSSVIRLRIVYQ 420
 QY 421 YPF 423
 DB 421 YPF 423
 RESULT 13
 AA56750
 ID AA56750 standard; Protein; 448 AA..
 XX
 XX AA56750;
 XX
 XX 22-FEB-2000 (first entry)
 XX
 XX Smooth muscle proliferation modulating protein.
 XX
 XX Cell proliferation; vascular; smooth muscle cell; arteriosclerosis; PTCA;
 XX endothelial thickening; percutaneous transluminal coronary angioplasty;
 XX myoma; hematopoietic cell-regulation; cytokine; tissue generation; taxis;
 XX actin activity; chemotaxis; blood coagulation; thrombotic; tumor; mouse;
 XX metastasis; nutrient.
 XX
 OS Mus musculus.
 XX
 XX WO9955863-A1.
 XX
 XX 04-NOV-1999.
 XX
 XX 28-APR-1999; 99WO-JP02283.
 XX
 XX 28-APR-1998; 98JP-0119731.
 XX
 XX (ONOX) ONO PHARM CO LTD.
 XX
 XX Honjo T, Tashiro K, Nakamura T;
 XX WPI: 2000-038646/03.
 XX N-PSDB; AA239383, AA239384.
 XX
 XX Novel mouse polypeptides for treatment, e.g. arteriosclerosis and myoma
 XX
 XX Claim 1; Page 42-44; 70pp; Japanese.
 XX
 CC The invention provides mouse polypeptides for treatment of diseases due
 CC to abnormal proliferation of smooth muscle. The polypeptides can be
 CC produced by standard recombinant methodology. The polypeptides can be
 CC used according to their inhibition of the proliferation of vascular
 CC smooth muscle cells, particularly in treating arteriosclerosis or re-
 CC narrowing by vascular endothelial thickening after percutaneous
 CC transluminal coronary angioplasty (PTCA), or myoma, hematopoietic cell-
 CC regulatory activity, cytokine activity, tissue generation/repatriation
 CC activity, actin/inhibitin activity, taxis and chemotaxis activity, blood
 CC coagulation/thrombotic activity, receptor/ligand activity, cadohelin/
 CC tumour metastasis inhibiting activity; tumor inhibition, and as nutrient.
 CC The present sequence represents the protein of the invention which can be
 CC used for modulating smooth muscle cell proliferation.
 CC
 XX Sequence 448 AA:
 SO
 Query Match 95.6%; Score 2302; DB 21; Length 448;
 Best Local Similarity 94.8%; Pred. No. 2.4e-153;
 Matches 401; Conservative 9; Mismatches 13; Indels 0; Gaps 0;
 QY 1 OCTNGFDLDRSGQCIDIDECPRIPEACRGDMACVNONGGYLCIPRTNRYRGPSNPS 60
 DB 26 OCTNGFDLDRSGQCIDIDECPRIPEACRGDMACVNONGGYLCIPRTNRYRGPSNPS 85

CC can be modulated with agents that bind to these polypeptides, resulting
CC in the death of the cells. The polynucleotides encoding these
CC polypeptides are useful in the recombinant production of the
CC polypeptides, as a hybridisation probe to screen libraries to isolate
CC homologous sequences, or to map the gene. They may also be used for
CC analysing genetic disorders, and to produce transgenic animals which are
CC useful for the development and screening of therapeutically useful
CC reagents. The polynucleotides can also be used in gene therapy e.g. to
CC replace a defective gene.

XX Sequence 448 AA;

Query Match 99.8%; Score 2401; DB 22; Length 448;
Best Local Similarity 99.8%; Pred. No. 2.8e-160;
Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 QCTNGFPLDRSGGCLDIDECRTPEACRGDMCMVNGGYLCIPRTNPPYRGPSNPS 60
DB 26 QCTNGFPLDRSGGCLDIDECRTPEACRGDMCMVNGGYLCIPRTNPPYRGPSNPS 85
QY 61 TPYSGPYPAAPPLSAPNYPITISRLICRGYOMDESOCVDDECATDSHCNPTQICI 120
DB 86 TPYSGPYPAAPPLSAPNYPITISRLICRGYOMDESOCVDDECATDSHCNPTQICI 145
QY 121 NTGEGYTCTDGYWLEGGCLDIDECRYGCOQLCANVPGSYSTCNPFTLNEDGRSC 180
DB 146 NTGEGYTCTDGYWLEGGCLDIDECRYGCOQLCANVPGSYSTCNPFTLNEDGRSC 205
QY 181 QDVNECATENPCVQTCVNTGSLCRDPGYELEEDGYHCSMDCSFSEFLCOHECVNQ 240
DB 206 QDVNECATENPCVQTCVNTGSLCRDPGYELEEDGYHCSMDCSFSEFLCOHECVNQ 265
QY 241 PGTYFSCSPGYILLDNRSCQDINECEHNRHTCNLOQTCYNLOGGKCIDPICEEPTL 300
DB 266 PGTYFSCSPGYILLDNRSCQDINECEHNRHTCNLOQTCYNLOGGKCIDPICEEPTL 325
QY 301 RISDNRCMCPAENPCGRDOPFTILYRDMDVSGRSVPADIFOMATTRYPGAYYIFQIKS 360
DB 326 RISDNRCMCPAENPCGRDOPFTILYRDMDVSGRSVPADIFOMATTRYPGAYYIFQIKS 385
QY 361 GNEGREFYRQGTGPISATLVMTTRPIKGPRIQDLEMTVNTVINFRSSVIRLRITYSQ 420
DB 386 GNEGREFYRQGTGPISATLVMTTRPIKGPRIQDLEMTVNTVINFRSSVIRLRITYSQ 445
QY 421 YPF 423
DB 446 YPF 448
b
```

RESULT 11
AA56751

ID AA56751 standard; Protein; 423 AA.

XX AA56751;

XX 22-FEB-2000 (first entry)

XX Smooth muscle proliferation modulating protein mature sequence.

XX Cell proliferation; vascular; smooth muscle cell; arteriosclerosis; PTCA;
XX endothelial thickening; percutaneous transluminal coronary angioplasty;
XX myoma; hematopoietic cell-regulation; cytokine; tissue generation; taxis;
XX actin activity; chemotaxis; blood coagulation; thrombotic; tumor; mouse;
XX metastasis; nutrient.

XX Mus musculus.

XX MO9955863-A1.

XX 04-NOV-1999.

XX 28-APR-1999; 99MO-JP02283.

PR 28-APR-1998; 98JP-0119731.
XX
XX (ONOI) ONO PHARM CO LTD.
XX
XX Honjo T, Tashiro K, Nakamura T;
XX
XX WPI; 2000-038646/03.
XX N-PSDB; AA239385.
XX
XX Novel mouse polypeptides for treatment, e.g. arteriosclerosis and myoma
XX
XX
XX Claim 1: Page 51-53; 70pp; Japanese.

The invention provides mouse polypeptides for treatment of diseases due
CC to abnormal proliferation of smooth muscle. The polypeptides can be
CC produced by standard recombinant methodology. The polypeptides can be
CC used according to their inhibition of the proliferation of vascular
CC smooth muscle cells, particularly in treating arteriosclerosis or re-
CC narrowing by vascular endothelial thickening after percutaneous
CC transluminal coronary angioplasty (PTCA), or myoma, hematopoietic cell-
CC regulatory activity, cytokine activity, tissue generation/regeneration
CC activity, actin/inhibitor activity, taxis and chemotaxis activity, blood
CC coagulation/thrombotic activity, receptor/ligand activity, cadherin/
CC tumour metastasis inhibiting activity; tumor inhibition, and as nutrient.
CC The present sequence represents the mature protein of the invention which
CC can be used for modulating smooth muscle cell proliferation.

Sequence 423 AA;

Query Match 95.6%; Score 2302; DB 21; Length 423;
Best Local Similarity 94.8%; Pred. No. 2.2e-153;
Matches 401; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

```
QY 1 QCTNGFPLDRSGGCLDIDECRTPEACRGDMCMVNGGYLCIPRTNPPYRGPSNPS 60
DB 1 QCTNGFPLDRSGGCLDIDECRTPEACRGDMCMVNGGYLCIPRTNPPYRGPSNPS 60
QY 61 TPYSGPYPAAPPLSAPNYPITISRLICRGYOMDESOCVDDECATDSHCNPTQICI 120
DB 61 TPYSGPYPAAPPLSAPNYPITISRLICRGYOMDESOCVDDECATDSHCNPTQICI 120
QY 121 NTGEGYTCTDGYWLEGGCLDIDECRYGCOQLCANVPGSYSTCNPFTLNEDGRSC 180
DB 121 NTGEGYTCTDGYWLEGGCLDIDECRYGCOQLCANVPGSYSTCNPFTLNEDGRSC 180
QY 181 QDVNECATENPCVQTCVNTGSLCRDPGYELEEDGYHCSMDCSFSEFLCOHECVNQ 240
DB 181 QDVNECATENPCVQTCVNTGSLCRDPGYELEEDGYHCSMDCSFSEFLCOHECVNQ 240
QY 241 PGTYFSCSPGYILLDNRSCQDINECEHNRHTCNLOQTCYNLOGGKCIDPICEEPTL 300
DB 241 PGTYFSCSPGYILLDNRSCQDINECEHNRHTCNLOQTCYNLOGGKCIDPICEEPTL 300
QY 301 RISDNRCMCPAENPCGRDOPFTILYRDMDVSGRSVPADIFOMATTRYPGAYYIFQIKS 360
DB 301 LISDNRCMCPAENPCGRDOPFTILYRDMDVSGRSVPADIFOMATTRYPGAYYIFQIKS 360
QY 361 GNEGREFYRQGTGPISATLVMTTRPIKGPRIQDLEMTVNTVINFRSSVIRLRITYSQ 420
DB 361 GNEGREFYRQGTGPISATLVMTTRPIKGPRIQDLEMTVNTVINFRSSVIRLRITYSQ 420
QY 421 YPF 423
DB 421 YPF 423
```

RESULT 12

AA56753

ID AA56753 standard; Protein; 423 AA.

XX AA56753;

XX

CC the proliferation or differentiation of chondrocyte cells. The PRO
CC proteins can be used to determine the presence of tumours and also
CC susceptibility to tumour development, particularly adenoma, lung, colon,
CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
CC can be used for genetic analysis of individuals with genetic disorders.
XX

Sequence 448 AA:

Query Match 99.8%; Score 2401; DB 22; Length 448;
Best Local Similarity 99.8%; Pred. No. 2,8e-160;
Matches 442; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 OCTNGFDLDRSGGCLDIDECRTIPEACRGDMCMVNGGYLCIPRTNPRYRGPSNPS 60
26 OCTNGFDLDRSGGCLDIDECRTIPEACRGDMCMVNGGYLCIPRTNPRYRGPSNPS 85
61 TPYSGPYPAAPLSPAPNPTISRPLICRGYQMDENQCVADDECATDSHOCNPTQICI 120
86 TPYSGPYPAAPLSPAPNPTISRPLICRGYQMDENQCVADDECATDSHOCNPTQICI 145
121 NTEGTYCTGCTDGYWLLGCGCLDIDECRYGCGQQLCANVPSSYSTCNCNPGFTLNDRGSC 180
146 NTEGTYCTGCTDGYWLLGCGCLDIDECRYGCGQQLCANVPSSYSTCNCNPGFTLNDRGSC 205
181 QDVNECATENPCVQTCVNTYGSFICRCDPGYELEDGVCSDMDECSFEFLQHECVNQ 240
206 QDVNECATENPCVQTCVNTYGSFICRCDPGYELEDGVCSDMDECSFEFLQHECVNQ 265
241 PGTYFSCPPGTYLLDNRSCQDINECEHRNHTCNLQOTCYNLGGFKCIPDIRCEEPYL 300
266 PGTYFSCPPGTYLLDNRSCQDINECEHRNHTCNLQOTCYNLGGFKCIPDIRCEEPYL 325
301 RISDNMCMCPAEMPGCNDOPFTLLYRDMVYSGRSVPADIDOMATTYPPAYYFQIKS 360
326 RISDNMCMCPAEMPGCNDOPFTLLYRDMVYSGRSVPADIDOMATTYPPAYYFQIKS 385
361 GNEGREFYKQTPISATLVMTPIKGPRIQDLEMTVTNFINRGSSVIRLRIYYSQ 420
386 GNEGREFYKQTPISATLVMTPIKGPRIQDLEMTVTNFINRGSSVIRLRIYYSQ 445
421 YPF 423
446 YPF 448

RESULT 10

ID AAB31183 standard; Protein; 448 AA.

XX AAB31183;

20-APR-2001 (first entry)

XX Amino acid sequence of human polypeptide PRO210.

XX Human; secreted protein; transmembrane protein; PRO196; PRO444; PRO183;
XX PRO185; PRO210; PRO215; PRO217; PRO242; PRO288; PRO365; PRO1361; PRO1308;
XX PRO1183; PRO1272; PRO1419; PRO4999; PRO1170; PRO248; PRO553; PRO1318;
XX PRO1600; PRO9940; PRO333; PRO301; PRO187; PRO337; PRO1411; PRO4356;
XX PRO246; PRO265; PRO941; PRO10096; PRO6003; PRO350; PRO2630;
XX PRO6309; cell death; genetic disorder; transgenic animal; gene therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Peptide 1..25 "signal peptide"

XX Modified-site 21..27 "N-myristoylation site"

XX Binding-site 54..57 "note- "cell attachment site"

XX Modified-site 64..70

FT /note- "N-myristoylation site"
FT 144..156
FT /note- "aspartic acid and asparagine hydroxylation site"
FT 149..155
FT /note- "N-myristoylation site"
FT 186..192
FT /note- "N-myristoylation site"
FT 226..232
FT /note- "N-myristoylation site"
FT 242..248
FT /note- "N-myristoylation site"
FT 267..273
FT /note- "N-myristoylation site"
FT 283..287
FT /note- "N-glycosylation site"
FT 296..300
FT /note- "N-glycosylation site"
FT 310..316
FT /note- "N-myristoylation site"

W0200077037-A2.

XX 21-DEC-2000.
XX 22-MAY-2000; 2000WO-US14042.
XX 15-JUN-1999; 99US-0139695.
XX 20-JUL-1999; 99US-0145070.
XX 26-JUL-1999; 99US-0145698.
XX 17-AUG-1999; 99US-0149396.
XX 01-SEP-1999; 99WO-US20111.
XX 08-SEP-1999; 99WO-US20594.
XX 15-SEP-1999; 99WO-US21090.
XX 15-SEP-1999; 99WO-US21547.
XX 30-NOV-1999; 99WO-US28313.
XX 01-DEC-1999; 99WO-US28301.
XX 02-DEC-1999; 99WO-US28565.
XX 07-DEC-1999; 99US-0169495.
XX 05-JAN-2000; 2000WO-US00219.
XX 18-FEB-2000; 2000WO-US04341.
XX 18-FEB-2000; 2000WO-US04342.
XX 22-FEB-2000; 2000WO-US04414.
XX 01-MAR-2000; 2000WO-US05601.
XX 02-MAR-2000; 2000WO-US05841.
XX 20-MAR-2000; 2000WO-US07377.
XX 30-MAR-2000; 2000WO-US08439.
XX 15-MAY-2000; 2000WO-US13358.
XX 17-MAY-2000; 2000WO-US13705.

(GETH) GENENTECH INC.

XX Ashkenazi AJ, Baker KP, Botstein DA, Desnoyers L, Eaton DL;
XX Ferrara N, Fong S, Gao W, Geider H, Gerritsen ME, Goddard A;
XX Godowski PJ, Gurney AL, Kljavin JV, Mather JP, Napier MA, Pan J;
XX Paoni NF, Roy MA, Stewart JA, Tumas D, Watanabe CK, Williams PW;
XX Wood WI, Zhang Z;

WPI: 2001-050091/06.

DR N-PSDB; AAC6968.

XX Isolated nucleic acid molecule encoding a PRO polypeptide which is a
XX transmembrane polypeptide is useful for gene therapy and identification
XX of related polypeptides -

PS Claim 12; Fig 10; 244pp; English.

XX The present sequence represents a human secreted and transmembrane
XX polypeptide. The specification describes human polypeptides, designated
XX PRO196, PRO444, PRO183, PRO210, PRO215, PRO217, PRO242, PRO288,
XX PRO365, PRO1361, PRO1308, PRO1183, PRO1272, PRO4999, PRO7170,
XX PRO248, PRO353, PRO1318, PRO1600, PRO9940, PRO533, PRO301, PRO187,
XX PRO337, PRO4111, PRO4356, PRO246, PRO265, PRO941, PRO10096, PRO6003,
XX PRO6004, PRO350, PRO2630 and PRO6309. The biological activity of cells

CC Hashimoto's thyroiditis, juvenile lymphocytic thyroiditis, atrophic
 CC thyroiditis), diabetes mellitus, immune-mediated renal disease
 CC (glomerulonephritis, tubulointerstitial nephritis), multiple sclerosis,
 CC idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome, chronic
 CC inflammatory demyelinating polyneuropathy, infectious hepatitis
 CC (hepatitis A, B, C, D, E and other non-hepatotropic viruses), autoimmune
 CC chronic active hepatitis, primary biliary cirrhosis, granulomatous
 CC hepatitis, and sclerosing cholangitis, inflammatory bowel disease
 CC (ulcerative colitis: Crohn's disease), gluten-sensitive enteropathy, and
 CC Whipple's disease. Autoimmune or immune-mediated skin diseases including
 CC bullous skin diseases, erythema multiforme, contact dermatitis, psoriasis,
 CC asthma, allergic rhinitis, atopic dermatitis, food hypersensitivity,
 CC urticaria, eosinophilic pneumonia, idiopathic pulmonary fibrosis,
 CC hypersensitivity pneumonitis, and transplantation associated diseases
 CC (graft rejection, and graft-versus-host-disease). (i) Its (ant)agonists
 CC or fragment can also be used as an adjuvant in treatment of tumors.
 CC Antibodies against (i) can also be used for diagnosing such diseases.
 CC This sequence represents a human EGF-like homologue encoded by cDNA clone
 CC DN32279 which is described in the invention.

XX Sequence 448 AA;

Query Match 99.8%; Score 2401; DB 20; Length 448;

Best Local Similarity 99.8%; Pred. No. 2.8e-160; Mismatches 1; Indels 0; Gaps 0;

Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 OCTNGFDLRSGGCLDIDECRTPEACRGDMCVNONGYLCIPRTNPPYRGPSNYS 60
 DB 26 OCTNGFDLRSGGCLDIDECRTPEACRGDMCVNONGYLCIPRTNPPYRGPSNYS 85
 QY 61 TPYSGPPAPAPPLSAPNYPYPTISPLICRGYQMDSESNQCVNDECATDSHCNPTQICI 120
 DB 86 TPYSGPPAPAPPLSAPNYPYPTISPLICRGYQMDSESNQCVNDECATDSHCNPTQICI 145
 QY 121 NTEGYTSCCTDGYWYLEGSCCLDIDECRYGCOQLCANVGSYSCCTNPGFTLNEDRSC 180
 DB 146 NTEGYTSCCTDGYWYLEGSCCLDIDECRYGCOQLCANVGSYSCCTNPGFTLNEDRSC 205
 QY 181 ODVNECATENPCVQTCVNTYGSFTCRCPGYELLEDGVHCSMDDECSFSEFLCHECVNQ 240
 DB 206 ODVNECATENPCVQTCVNTYGSFTCRCPGYELLEDGVHCSMDDECSFSEFLCHECVNQ 265
 QY 241 PGTYFSCPPGYLLDDNSCODINCEHNRHNCNLOQTCYNLOGGKCDIPRCEPYL 300
 DB 266 PGTYFSCPPGYLLDDNSCODINCEHNRHNCNLOQTCYNLOGGKCDIPRCEPYL 325
 Y 301 RLSDRNCMPAENPCGRDOPFTLLYRDMOVSGRSVPADIFQMATRYPGAYYIFQIKS 360
 DB 326 RLSDRNCMPAENPCGRDOPFTLLYRDMOVSGRSVPADIFQMATRYPGAYYIFQIKS 385
 QY 361 GNEGREFFYRGTGPISATLVMTRPITKGPRIQDLEMTIVNTYINRGSSVILRLYVSQ 420
 DB 386 GNEGREFFYRGTGPISATLVMTRPITKGPRIQDLEMTIVNTYINRGSSVILRLYVSQ 445
 QY 421 YPF 423
 DB 446 YPF 448

RESULT 9

AAU29227 standard: Protein; 448 AA.

AAU29227;

18-DEC-2001 (first entry)

Human PRO polypeptide sequence #204.

XX PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
 KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
 KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
 KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.

XX OS Homo sapiens.
 XX PN WO200168848-A2.
 XX PD 20-SEP-2001.
 XX PF 28-FEB-2001; 2001WO-US06520.
 XX PR 01-MAR-2000; 2000WO-US05601.
 XX PR 02-MAR-2000; 2000WO-US05841.
 XX PR 03-MAR-2000; 2000US-187202P.
 XX PR 06-MAR-2000; 2000US-186968P.
 XX PR 14-MAR-2000; 2000US-189320P.
 XX PR 14-MAR-2000; 2000US-189328P.
 XX PR 15-MAR-2000; 2000WO-US06884.
 XX PR 21-MAR-2000; 2000US-190828P.
 XX PR 21-MAR-2000; 2000US-191007P.
 XX PR 21-MAR-2000; 2000US-191048P.
 XX PR 21-MAR-2000; 2000US-191314P.
 XX PR 28-MAR-2000; 2000US-192655P.
 XX PR 29-MAR-2000; 2000US-193032P.
 XX PR 29-MAR-2000; 2000US-193053P.
 XX PR 30-MAR-2000; 2000WO-US08439.
 XX PR 04-APR-2000; 2000US-194449P.
 XX PR 04-APR-2000; 2000US-194647P.
 XX PR 11-APR-2000; 2000US-195975P.
 XX PR 11-APR-2000; 2000US-196000P.
 XX PR 11-APR-2000; 2000US-196187P.
 XX PR 11-APR-2000; 2000US-196690P.
 XX PR 11-APR-2000; 2000US-196820P.
 XX PR 18-APR-2000; 2000US-198121P.
 XX PR 18-APR-2000; 2000US-198585P.
 XX PR 25-APR-2000; 2000US-199397P.
 XX PR 25-APR-2000; 2000US-199550P.
 XX PR 03-MAY-2000; 2000US-201516P.
 XX PR 17-MAY-2000; 2000WO-US13705.
 XX PR 22-MAY-2000; 2000WO-US14042.
 XX PR 30-MAY-2000; 2000WO-US14941.
 XX PR 02-JUN-2000; 2000WO-US15264.
 XX PR 05-JUN-2000; 2000US-209832P.
 XX PR 28-JUL-2000; 2000WO-US20710.
 XX PR 22-AUG-2000; 2000US-064484P.
 XX PR 24-AUG-2000; 2000WO-US23328.
 XX PR 08-NOV-2000; 2000WO-US30952.
 XX PR 01-DEC-2000; 2000WO-US32678.
 XX PR 20-DEC-2000; 2000WO-US34956.
 XX (GENTH) GENENTECH INC.
 XX PA Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
 XX PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
 XX PI WPI; 2001-602746/68.
 XX DR N-PSDB; AAS46128.
 XX PT Novel nucleic acids encoding PRO polypeptides, used to diagnose the
 XX PT presence of tumours, such as prostate and breast tumours, in mammals and
 XX PT to screen for modulators of the compounds -
 XX Claim 11; Fig 408; 774pp; English.
 XX Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
 XX The PRO polypeptides and their associated nucleic acids can be used to
 XX detect the presence of a tumour in a mammal by comparing the level of
 XX expression of a PRO polypeptide in a test sample of cells from the animal
 XX and a control sample of normal cells, whereby a higher level of
 XX expression in the test sample indicates the presence of a tumour in the
 XX mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
 XX and rabbits but are preferably human. The polypeptides can be used to
 XX stimulate tumour necrosis factor (TNF) alpha release from human blood,
 XX when contacted with it. A specific polypeptide can be used to stimulate

PN US001051358-A1.
 XX
 PD 13-DEC-2001.
 XX
 PF 25-MAR-1999; 99US-0275805.
 XX
 PR 11-APR-1997; 97US-0839525.
 PR 10-APR-1996; 96WO-US05247.
 XX
 PA (OLSEN/) OLSEN H S.
 PA (LIHH/) LI H.
 PI Olsen HS, LI H;
 DR WPI: 2002-121417/16.
 DR N-PSDB: ABRK3627.
 XX
 PT New nucleic acid encoding human extracellular/epidermal growth factor,
 PT useful for diagnosis and treatment of e.g. Marfan syndrome and wounds,
 PT also related polypeptides -
 XX
 PS Claim 10; Fig 1; 22pp; English.
 CC The invention relates to a novel polynucleotide which is at
 CC least 95% identical with a sequence (ATCC 97285) encoding mature human
 CC extracellular protein-like/epidermal growth factor (BGF)-like protein,
 CC EGF. Also included are the EGF EGF domains, a vector containing
 CC the polynucleotide, a host cell containing the vector, anti-EGF
 CC antibodies and antagonists of EGF. The polynucleotide is used for
 CC recombinant production of EGF, in gene therapy, as hybridisation probes,
 CC as antisense antagonists and for chromosome identification. The protein
 CC is used to treat patients who require EGF, to identify specific
 CC antagonists, used to treat conditions that require inhibition of EGF
 CC (e.g. vascular smooth muscle cell proliferation, Marfan syndrome, wound
 CC healing, neurological trauma, acquired immunodeficiency syndrome,
 CC AIDS)-related dementia, ocular disorders, kidney disorders, liver
 CC disorders, hair follicle growth promotion, burns, ulcers, corneal
 CC incisions, corneal inflammation, neoplasms and psoriasis), to raise
 CC specific antibodies and to characterise receptors. The present
 CC sequence represents EGF.
 CC
 SQ Sequence 448 AA:
 Query Match 100.0%; Score 2407; DB 23; Length 448;
 Best Local Similarity 100.0%; Pred. No. 1.1e-160;
 Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OCTNGDLDROSGCCLDIDECRTIPEACRGDMCVNONGYLCIPRNPVTRGPTSNPYS 60
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 26 OCTNGDLDROSGCCLDIDECRTIPEACRGDMCVNONGYLCIPRNPVTRGPTSNPYS 85
 QY 61 TPYSGPYPAAPPLSAPNPTISNPLICRGYOMDESNOCDVDEACADSHQCNPTQICI 120
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 86 TPYSGPYPAAPPLSAPNPTISNPLICRGYOMDESNOCDVDEACADSHQCNPTQICI 145
 QY 121 NTEGGYTCTDGYWLEGGCLDIDECRYGYCOQLCANVPYSYCTCNPGFTLNEDGRSC 180
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 146 NTEGGYTCTDGYWLEGGCLDIDECRYGYCOQLCANVPYSYCTCNPGFTLNEDGRSC 205
 QY 181 QDVNECATENPCVOTCVNTGSGFICRDPGYELEEDGYHGDMDSCSSEFLCOHECYNQ 240
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 206 QDVNECATENPCVOTCVNTGSGFICRDPGYELEEDGYHGDMDSCSSEFLCOHECYNQ 265
 QY 241 PCTYFSCSPGYTLLDNRSCODINECEHNRHTCNLOOTCVNLGGFCIDPIRCEEPYL 300
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 266 PCTYFSCSPGYTLLDNRSCODINECEHNRHTCNLOOTCVNLGGFCIDPIRCEEPYL 325
 QY 301 RISDNKCMCPAENPGCRDOPFTLLYRDMVYSGRSVPADIFQOMATTTPGAYYIFQJKS 360
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 326 RISDNKCMCPAENPGCRDOPFTLLYRDMVYSGRSVPADIFQOMATTTPGAYYIFQJKS 385
 QY 361 GNEGREFYKQRTGPISATLVMTRPRIKPREIQLDLEMITVNTVINFGSSVIRLRIVYSQ 420
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 386 GNEGREFYKQRTGPISATLVMTRPRIKPREIQLDLEMITVNTVINFGSSVIRLRIVYSQ 445
 QY 421 YPF 423
 DB |||
 DB 446 YPF 448
 RESULT 8
 AAY08063
 ID AAY08063 standard; Protein: 448 AA.
 AC AAY08063;
 XX
 DT 11-SEP-2000 (first entry)
 XX
 DE Human EGF-like homologue protein (PRO217) encoded by DNA32279 cDNA.
 XX
 KW Inflammatory cell infiltration; immune response; T cell proliferation;
 KW anti-inflammatory; anti-autoimmune; anti-diabetic; spondyloarthropathy;
 KW T cell-mediated disease; spondyloarthropathy; sclerosis; renal disease;
 KW inflammatory myopathy; hemolytic anemia; thrombocytopenia; thyroiditis;
 KW diabetes mellitus; demyelinating polynuropathy; Guillain-Barre syndrome;
 KW multiple sclerosis; polynuropathy; hepatitis; cirrhosis; enteropathy;
 KW sclerosing cholangitis; inflammatory bowel disease; Whipple's disease;
 KW skin disease; dermatitis; psoriasis; asthma; allergic rhinitis; tumor;
 KW food hypersensitivity; urticaria; eosinophilic pneumonia; transplant;
 KW idiopathic pulmonary fibrosis; graft rejection; PRO245; human; PRO217;
 KW EGF-like.
 XX
 OS Homo sapiens.
 ON
 PN WO9914241-A2.
 XX
 PD 25-MAR-1999.
 XX
 PE 17-SEP-1998; 98WO-US19437.
 PF
 PR 17-SEP-1997; 97US-0059119.
 PR 18-SEP-1997; 97US-0059263.
 PR 28-OCT-1997; 97US-0063550.
 PR 12-NOV-1997; 97US-0065186.
 PR 21-NOV-1997; 97US-0066364.
 PR 24-NOV-1997; 97US-0066770.
 PR 04-JUN-1998; 98US-0088026.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Fong S, Goddard A, Gurney AL, Tumas D, Wood WT;
 XX
 DR WPI: 1999-229499/19.
 DR N-PSDB: AAX37670.
 XX
 PT Composition containing novel polypeptide PRO245, its agonist or
 PT antagonist -
 XX
 PS Example 1; Fig 6A; 177pp; English.
 XX
 CC This invention describes a novel composition containing (apart from a
 CC carrier or excipient) a novel PRO245 polypeptide (II). Its agonist or
 CC antagonist, or their fragments, for modulating: (i) Infiltration of
 CC inflammatory cells into tissue; (ii) an immune response; or (iii) T cell
 CC proliferation. The composition increases or decreases any of the effects
 CC (i)-(iii). The products of the invention have anti-inflammatory,
 CC anti-autoimmune and anti-diabetic activity. (i), and its (anti)agonists
 CC and their fragments, are used to treat immune-related diseases,
 CC particularly T cell-mediated diseases. The diseases treated include
 CC systemic lupus erythematosus, rheumatoid arthritis, juvenile chronic
 CC arthritis, spondyloarthropathies, systemic sclerosis (scleroderma),
 CC idiopathic inflammatory myopathies (dermatomyositis, polymyositis),
 CC Sjogren's syndrome, systemic vasculitis, sarcoidosis, autoimmune
 CC hemolytic anemia (immune pancytopenia, paroxysmal nocturnal
 CC hemoglobinuria), autoimmune thrombocytopenia (idiopathic thrombocytopenic
 CC purpura immune-mediated thrombocytopenia), thyroiditis (grave's disease,

Db 266 PGTYFCSPGXYILLDDNRSCODINECEHNRHTCNLOQTCVNLGGFKCIDPRICEPYL 325
 QY 301 RISDNRCMCPAENPCGRDOPFTILYRDMDVSGRSVPADIFQMOATRRYPGAYYIFQIKS 360
 Db 326 RISDNRCMCPAENPCGRDOPFTILYRDMDVSGRSVPADIFQMOATRRYPGAYYIFQIKS 385
 QY 361 GNEGREFYMRQTGPISATLVMTRPKIGPREIOLDLEMTVNTVINFRGSSVIRLRIVSQ 420
 Db 386 GNEGREFYMRQTGPISATLVMTRPKIGPREIOLDLEMTVNTVINFRGSSVIRLRIVSQ 445
 QY 421 YPF 423
 Db 446 YPF 448

RESULT 6
 AAM93573
 ID AAM93573 standard; Protein; 448 AA.

XX AC AAM93573;
 XX DT 06-NOV-2001 (first entry)
 XX DE Human polypeptide, SEQ ID NO: 3357.
 XX KW Human; full length cDNA; cDNA synthesis; oligo-capping.
 XX OS Homo sapiens.
 XX PN EP1130094-A2.
 XX PD 05-SEP-2001.
 XX PF 07-JUL-2000; 2000EP-0114089.
 XX PR 08-JUL-1999; 99JP-0194486.
 XX PR 11-JAN-2000; 2000JP-0118774.
 XX PR 02-MAY-2000; 2000JP-0183765.
 XX PA (HELI-) HELIX RES INST.
 XX PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 XX PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX DR WPI: 2001-524255/58.
 XX DR N-PSDB: AAK94505.
 XX PT 830 Primers useful for synthesizing full length cDNA clones and their
 XX PT use in genetic manipulation -

PS Claim 8: SEQ ID NO 3357; 1380bp + sequence listing; English.
 XX CC The invention relates to primers for synthesizing full length cDNA
 XX CC clones. 830 cDNA molecules encoding a human protein have been
 XX CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 XX CC molecules have been determined. Primers for synthesizing the full length
 XX CC cDNA are useful for clarifying the function of the protein encoded by
 XX CC the cDNA. The full length clones were obtained by construction of full
 XX CC length enriched cDNA libraries that were synthesized by the oligo-capping
 XX CC method. The primers enable the production of the full length cDNA easily
 XX CC without any special methods. The present sequence is a polypeptide
 XX CC encoded by a full length human cDNA of the invention.
 XX CC Note: The sequence data for this patent did not form part of the printed
 XX CC specification, but was obtained in CD-ROM format directly from EPO.
 XX SQ Sequence 448 AA:

Query Match 100.0%; Score 2407; DB 22; Length 448;
 Best Local Similarity 100.0%; Pred. No. 1.le-160;
 Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QCTNGFDLDRSGGCLDIDECRTIPACRGDMCVNONGYLICIPRTNPPYRGPSNPYS 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 26 PGTYFCSPGXYILLDDNRSCODINECEHNRHTCNLOQTCVNLGGFKCIDPRICEPYL 300
 QY 61 TPYSGPYPAAPPLSAPNYPILSRPLICRFYQMDENOCVDNDECADSHQCNTOICI 120
 Db 86 TPYSGPYPAAPPLSAPNYPILSRPLICRFYQMDENOCVDNDECADSHQCNTOICI 145
 QY 121 NTEGGYTCSTGDMLEGGCLDIDECRYGCOQLCANVPSSCTCNPGFTLNEDGRSC 180
 Db 146 NTEGGYTCSTGDMLEGGCLDIDECRYGCOQLCANVPSSCTCNPGFTLNEDGRSC 205
 QY 181 ODVNECATENPCVQTCVNTGSGFICRCDDPGYLEEDGVHCSMDCECSFEFLCOHECVNQ 240
 Db 206 ODVNECATENPCVQTCVNTGSGFICRCDDPGYLEEDGVHCSMDCECSFEFLCOHECVNQ 265
 QY 241 PGTYFCSPGXYILLDDNRSCODINECEHNRHTCNLOQTCVNLGGFKCIDPRICEPYL 300
 Db 266 PGTYFCSPGXYILLDDNRSCODINECEHNRHTCNLOQTCVNLGGFKCIDPRICEPYL 325
 QY 301 RISDNRCMCPAENPCGRDOPFTILYRDMDVSGRSVPADIFQMOATRRYPGAYYIFQIKS 360
 Db 326 RISDNRCMCPAENPCGRDOPFTILYRDMDVSGRSVPADIFQMOATRRYPGAYYIFQIKS 385
 QY 361 GNEGREFYMRQTGPISATLVMTRPKIGPREIOLDLEMTVNTVINFRGSSVIRLRIVSQ 420
 Db 386 GNEGREFYMRQTGPISATLVMTRPKIGPREIOLDLEMTVNTVINFRGSSVIRLRIVSQ 445
 QY 421 YPF 423
 Db 446 YPF 448

RESULT 7
 AAU75494
 ID AAU75494 standard; Protein; 448 AA.

XX AC AAU75494;
 XX DT 23-APR-2002 (first entry)
 XX DE Human extracellular protein-1like/EGF-1like protein, EEGF.
 XX KW Human; extracellular protein-1like protein; EGF-1like;
 XX KW protein; epidermal growth factor; EEGF. ATCC 97285; gene therapy;
 XX KW vascular smooth muscle cell proliferation; Marfan syndrome;
 XX KW wound healing; neurological trauma; acquired immunodeficiency syndrome;
 XX KW AIDS-related dementia; ocular disorder; kidney disorder; liver disorder;
 XX KW hair follicle growth promotion; burn; ulcer; corneal incision;
 XX KW corneal inflammation; neoplasm; psoriasis.
 XX OS Homo sapiens.

XX FH Key
 XX FH Location/Qualifiers
 XX FT 1..25
 XX FT /label= Signal_peptide
 XX FT 26..448
 XX FT /label= Mature_EEGF
 XX FT /note= "This region is specifically claimed in claim 10"
 XX FT 112..153
 XX FT /label= EGF_1_domain
 XX FT /note= "This domain is specifically claimed in claim 10"
 XX FT 154..190
 XX FT /label= EGF_2_domain
 XX FT /note= "This domain is specifically claimed in claim 10"
 XX FT 191..230
 XX FT /label= EGF_3_domain
 XX FT /note= "This domain is specifically claimed in claim 10"
 XX FT 231..271
 XX FT /label= EGF_4_domain
 XX FT /note= "This domain is specifically claimed in claim 10"
 XX FT 272..314
 XX FT /label= EGF_5_domain
 XX FT /note= "This domain is specifically claimed in claim 10"

PT smooth muscle cell growth, vasculogenesis, restenosis or
XX atherosclerosis
PS Claim 4: Page 46-47: 49pp: English.

XX This is the amino acid sequence of the human secreted protein AK647. The
CC polynucleotide sequence was obtained from a human foetal kidney cDNA
CC library. AK647 homologues in chicks and rodents are involved in aortic
CC tissue development. The spatial and temporal distribution of AK647
CC indicated that it acts as an a modulator of smooth muscle cells in
CC vasculogenesis during embryonic development. The primary structure of
CC AK647 consists of multiple EGF domains. The AK647 protein can be used as
CC a nutritional source or supplement. The protein shows both inhibitory and
CC inducing, cytokine, cell proliferation and cell differentiation activity.
CC The protein may also be used in the treatment of immune deficiencies and
CC disorders, including severe combined immunodeficiency (SCID), HIV and
CC other viral, bacterial and fungal infections. Regulation of immune
CC responses may also be carried out by the AK647 protein. Other uses of the
CC protein include a role in the regulation of haematopoiesis and in the
CC treatment of myeloid and lymphoid cell deficiencies. Uses in bone,
CC cartilage, tendon, ligament and nerve tissue regrowth are also possible,
CC as well as for wound healing and in the treatment of ulcers and burns.
CC The polynucleotides and proteins can be used for preventing, treating or
CC ameliorating smooth muscle cell growth, vasculogenesis, restenosis,
CC atherosclerosis, blood vessel remodelling and degeneration. The proteins
CC may also have activin/inhibin, chemotactic/chemokinetic, haemostatic and
CC thrombolytic, receptor/ligand, anti-inflammatory, cadherin/tumour
CC invasion suppressor, and tumour inhibition activity. AK647 specific
CC antibodies can be used for promoting smooth muscle cell growth or
CC vasculogenesis. The proteins and polynucleotides can also be used for
CC detection, diagnosis and drug screening.

SQ Sequence 448 AA:

Query Match 100.0%; Score 2407; DB 21; Length 448;

Best Local Similarity 100.0%; Pred. No. 1.1e-160; Indels 0; Gaps 0;

Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OCTNGEFLDRLDSGCLDIDECRTIPEACRGDMCVNNGGYLCIRTPNVYRGPSNPS 60
DB 26 OCTNGEFLDRLDSGCLDIDECRTIPEACRGDMCVNNGGYLCIRTPNVYRGPSNPS 85
QY 61 TPYSGPYPAAPPLSAPNPTISRPILCRFGYQMDENOCVDVDECATDSHOCNPTQICI 120
DB 86 TPYSGPYPAAPPLSAPNPTISRPILCRFGYQMDENOCVDVDECATDSHOCNPTQICI 145
QY 121 NTGEGYTCTGCTDGYWLEGGCLDIDECRYGCGQQLCAVNPSSYSTCNPNGFTLNEDGRSC 180
DB 146 NTGEGYTCTGCTDGYWLEGGCLDIDECRYGCGQQLCAVNPSSYSTCNPNGFTLNEDGRSC 205
QY 181 QDVNECATENPCVOTCVNTYGSFICRCDPGYLEEDGYHCSMDSCSFSEFLCOHECVNQ 240
DB 206 QDVNECATENPCVOTCVNTYGSFICRCDPGYLEEDGYHCSMDSCSFSEFLCOHECVNQ 265
QY 241 PGTYFSCPPGYIILLDNRSQDINECEHRNHTCNLQOTCYNLQGGFKCIDIRCEEPYL 300
DB 266 PGTYFSCPPGYIILLDNRSQDINECEHRNHTCNLQOTCYNLQGGFKCIDIRCEEPYL 325
QY 301 RISDNRCMCAENPGCDOPFTILYRDMVYSGSVPADIFOMATTRYPGAYYIFQIKS 360
DB 326 RISDNRCMCAENPGCDOPFTILYRDMVYSGSVPADIFOMATTRYPGAYYIFQIKS 385
QY 361 GNGRGREYMQTGPISATLVTRPIKGPRIQDLDEMIVTVINFRGSSVRLRLYYSQ 420
DB 386 GNGRGREYMQTGPISATLVTRPIKGPRIQDLDEMIVTVINFRGSSVRLRLYYSQ 445
QY 421 YPF 423
DB 446 YPF 448

RESULT 5
AAV54989

ID AAV54989 standard; Protein: 448 AA.

XX AAV54989;

XX 15-FEB-2000 (first entry)

DE Full length human A55 protein sequence.

XX A55 protein: human; smooth muscle proliferation; tissue generation;
KW vascular smooth muscle cell; arteriosclerosis; tissue repair; myoma;
KW vascular endothelial thickening; haematopoietic cell regulator; cytokine;
KW percutaneous transluminal coronary angioplasty; blood coagulation; PTCA;
KW inhibin; chemotaxis; thrombosis; cadherin; therapy;
KW tumour metastasis inhibitor.

OS Homo sapiens.

PN W0955864-A1.

PD 04-NOV-1999.

XX 28-APR-1999; 99WC-JP02284.

XX 28-APR-1998; 98JP-0119731.

PA (ONOX) ONO PHARM CO LTD.

PI Honjo T, Tashiro K, Nakamura T;

DR WPI: 2000-038647/03.

DR N-PSDB; AA240027.

PT Novel human polypeptides for treatment of, e.g. arteriosclerosis and

XX myoma

PS Claim 1: Page 76-80, 87pp: Japanese.

XX This sequence is the human A55 protein of the invention. The protein
CC can be used for the treatment of diseases due to abnormal proliferation
CC of smooth muscle. The polypeptides can be used according to their inhibition
CC of the proliferation of vascular smooth muscle cells, particularly in
CC treating arteriosclerosis or re-narrowing by vascular endothelial
CC thickening after percutaneous transluminal coronary angioplasty (PTCA),
CC or myoma, haematopoietic cell-regulatory activity, cytokine activity,
CC tissue generation/repairation activity, actin/inhibin activity, taxis
CC and chemotaxis activity, blood coagulation/thrombotic activity,
CC receptor/ligand activity, cadherin/tumour metastasis inhibiting activity;
CC tumour inhibition, and as a nutrient.

SQ Sequence 448 AA:

Query Match 100.0%; Score 2407; DB 21; Length 448;

Best Local Similarity 100.0%; Pred. No. 1.1e-160; Indels 0; Gaps 0;

Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OCTNGEFLDRLDSGCLDIDECRTIPEACRGDMCVNNGGYLCIRTPNVYRGPSNPS 60
DB 26 OCTNGEFLDRLDSGCLDIDECRTIPEACRGDMCVNNGGYLCIRTPNVYRGPSNPS 85
QY 61 TPYSGPYPAAPPLSAPNPTISRPILCRFGYQMDENOCVDVDECATDSHOCNPTQICI 120
DB 86 TPYSGPYPAAPPLSAPNPTISRPILCRFGYQMDENOCVDVDECATDSHOCNPTQICI 145
QY 121 NTGEGYTCTGCTDGYWLEGGCLDIDECRYGCGQQLCAVNPSSYSTCNPNGFTLNEDGRSC 180
DB 146 NTGEGYTCTGCTDGYWLEGGCLDIDECRYGCGQQLCAVNPSSYSTCNPNGFTLNEDGRSC 205
QY 181 QDVNECATENPCVOTCVNTYGSFICRCDPGYLEEDGYHCSMDSCSFSEFLCOHECVNQ 240
DB 206 QDVNECATENPCVOTCVNTYGSFICRCDPGYLEEDGYHCSMDSCSFSEFLCOHECVNQ 265
QY 241 PGTYFSCPPGYIILLDNRSQDINECEHRNHTCNLQOTCYNLQGGFKCIDIRCEEPYL 300

```

QY 61 TPYSGPYPAAPPLSAPNPTISRPILCRFGYQMDSESNOCVDVDECATDSHCNPQICI 120
DB 86 TPYSGPYPAAPPLSAPNPTISRPILCRFGYQMDSESNOCVDVDECATDSHCNPQICI 145
QY 121 NTEGGYTCSTGTYWLLLEGQCLDIDECRYGYCOQLCANVPGSYSCNPGFTLNEDGRSC 180
DB 146 NTEGGYTCSTGTYWLLLEGQCLDIDECRYGYCOQLCANVPGSYSCNPGFTLNEDGRSC 205
QY 181 QDVNECATENPCVOTCVNTYGSFICRCDDPGYELEDVGHCSMDDECSFSEFLCOHECVNQ 240
DB 206 QDVNECATENPCVOTCVNTYGSFICRCDDPGYELEDVGHCSMDDECSFSEFLCOHECVNQ 265
QY 241 PGTYFSCSPPGYIILLDDNRSQCODINECEHNNHTCNLOQTCYNLOGGFKCIDPICEPYL 300
DB 266 PGTYFSCSPPGYIILLDDNRSQCODINECEHNNHTCNLOQTCYNLOGGFKCIDPICEPYL 325
QY 301 RISDNRCMCPAENPGCRDQPFITILYRDMDVYSGRSVPADIFQMATRRPGAYYIFQIKS 360
DB 326 RISDNRCMCPAENPGCRDQPFITILYRDMDVYSGRSVPADIFQMATRRPGAYYIFQIKS 385
QY 361 GNEGREFYMRQTPISATLVMTPIKGPREDIQLDEMTVNTVINFRGSSVIRLRIVSQ 420
DB 386 GNEGREFYMRQTPISATLVMTPIKGPREDIQLDEMTVNTVINFRGSSVIRLRIVSQ 445
QY 421 YPF 423
DB 446 YPF 448

RESULT 3
AAW94281
ID AAW94281 standard; Protein: 448 AA.
AC AAW94281;
XX 07-MAY-1999 (first entry)
XX Human extracellular matrix protein (ECMP)-1.
DE Extracellular matrix protein; ECMP-1; diagnosis; treatment; cancer;
KW Immune disorder; human.
XX Homo sapiens.
XX OS
XX PN W09900410-A2.
XX PD 07-JAN-1999.
XX PF 23-JUN-1998; 98MO-US13012.
XX PR 27-JUN-1997; 97US-0884072.
XX PA (INCY-) INCYTE PHARM INC.
XX PI Bandman O, Corley NC, Guegler KJ;
XX DR MPI: 1999-095674/08.
XX DR N-PSDB: AAX05359.
XX PT New polynucleotide encoding extracellular matrix protein, ECMP-1 -
XX PT disorders and cancer
XX PS Claim 1; Fig 1A-G; 79pp; English.
XX CC This represents a human extracellular matrix protein (ECMP)-1. Host
XX CC cells containing a vector comprising the ECMP-1 nucleic acid are used
XX CC for the recombinant production of the protein. ECMP-1 and its
XX CC (ant)agonists, are useful in the diagnosis, prevention, and treatment
XX CC of cancer and immune disorders.
XX SQ Sequence 448 AA;

```

Query Match 100.0%; Score 2407; DB 20; Length 448;
 Best Local Similarity 100.0%; Pred. No. 1.1e-160;
 Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 OCTGEPFLDROSGQCLDIDECRTIPEACRGCMCVNONGYLICIPRTNPVYRGPSNPYS 60
DB 26 QCTNGEPFLDROSGQCLDIDECRTIPEACRGCMCVNONGYLICIPRTNPVYRGPSNPYS 85
QY 61 TPYSGPYPAAPPLSAPNPTISRPILCRFGYQMDSESNOCVDVDECATDSHCNPQICI 120
DB 86 TPYSGPYPAAPPLSAPNPTISRPILCRFGYQMDSESNOCVDVDECATDSHCNPQICI 145
QY 121 NTEGGYTCSTGTYWLLLEGQCLDIDECRYGYCOQLCANVPGSYSCNPGFTLNEDGRSC 180
DB 146 NTEGGYTCSTGTYWLLLEGQCLDIDECRYGYCOQLCANVPGSYSCNPGFTLNEDGRSC 205
QY 181 QDVNECATENPCVOTCVNTYGSFICRCDDPGYELEDVGHCSMDDECSFSEFLCOHECVNQ 240
DB 206 QDVNECATENPCVOTCVNTYGSFICRCDDPGYELEDVGHCSMDDECSFSEFLCOHECVNQ 265
QY 241 PGTYFSCSPPGYIILLDDNRSQCODINECEHNNHTCNLOQTCYNLOGGFKCIDPICEPYL 300
DB 266 PGTYFSCSPPGYIILLDDNRSQCODINECEHNNHTCNLOQTCYNLOGGFKCIDPICEPYL 325
QY 301 RISDNRCMCPAENPGCRDQPFITILYRDMDVYSGRSVPADIFQMATRRPGAYYIFQIKS 360
DB 326 RISDNRCMCPAENPGCRDQPFITILYRDMDVYSGRSVPADIFQMATRRPGAYYIFQIKS 385
QY 361 GNEGREFYMRQTPISATLVMTPIKGPREDIQLDEMTVNTVINFRGSSVIRLRIVSQ 420
DB 386 GNEGREFYMRQTPISATLVMTPIKGPREDIQLDEMTVNTVINFRGSSVIRLRIVSQ 445
QY 421 YPF 423
DB 446 YPF 448

```

```

RESULT 4
AAV57058
ID AAV57058 standard; Protein: 448 AA.
AC AAV57058;
XX 21-FEB-2000 (first entry)
XX Amino acid sequence of the human secreted protein AK647.
DE AK647; aortic tissue development; smooth muscle cell modulator; SCIP;
KW nutritional supplement; vasculogenesis; embryonic development; infection;
KW cytokine activity; cell proliferation; cell differentiation; detect; HIV;
KW immune deficiency; haematopoiesis regulation; tissue regrowth; diagnose;
KW wound healing; restenosis; atherosclerosis; drug screen.
XX Homo sapiens.
XX OS
XX PN W09960125-A2.
XX PD 25-NOV-1999.
XX PF 18-MAY-1999; 99MO-US10931.
XX PR 19-MAY-1998; 98US-0081002.
XX PR 21-MAY-1998; 98US-0083002.
XX PA (GENY ) GENETICS INST INC.
XX PI Jacobs K, McCoy JM, Racie L, Lavallie E, Treacy M, Evans C;
XX PI Agostino M, Lu Z, Merberg D;
XX DR MPI: 2000-053298/04.
XX DR N-PSDB: AAZ39892.
XX PT Proteins, and their encoding polynucleotides, used for treating e.g.

```

FT Region /label= EGF-4
 FT 272..314
 FT /label= EGF-5
 XX
 PN MO846746-A1.
 XX
 PD 22-OCT-1998.
 XX
 PF 11-APR-1997; 97MO-US06020.
 XX
 PK 11-APR-1997; 97MO-US06020.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI LI H, Olsen HS;
 XX
 DR WPI; 1998-568728/48.
 XX
 DR N-PSDB; AAV62432.
 XX
 PT New isolated extracellular/epidermal growth factor - used for
 regulating vascular smooth muscle cell proliferation, e.g. for
 enhancing neurological functions or treating neoplasia and other
 disorders.

PS Claim 10a; Fig 1A-D; 62pp; English.

XX This sequence represents a novel human extracellular/epidermal growth
 CC factor-like protein, EGF. This protein can be used to regulate
 CC vascular smooth muscle cell proliferation and for restoration or
 CC enhancement of neurological functions diminished as a result of other
 CC damaging pathologies such as AIDS dementia. The protein can also be used
 CC to treat senile dementia, ocular disorders such as corneal inflammation,
 CC for targeting tumour cells, for treating kidney disorders, for liver
 CC regeneration or treating liver dysfunction, for treating wounds including
 CC all cutaneous wounds, corneal wounds, and injuries to the
 CC epithelial-lined hollow organs of the body or resulting from trauma such
 CC as burns, abrasions and cuts as well as from surgical procedures such as
 CC surgical incisions and skin grafting. The polypeptides can also be used
 CC for treating chronic conditions, such as chronic ulcers, diabetic ulcers,
 CC other non-healing (trophic) conditions, to treat Marfan syndrome, to
 CC promote hair follicular development, to stimulate growth and
 CC differentiation of various epidermal and epithelial cells in vivo and in
 CC vitro and to stimulate embryogenesis. Antagonists to EGF can be used to
 CC treat neoplasia such as cancers or tumours, skin disorders such as
 CC psoriasis or corneal inflammation. The products can also be used for
 CC identifying EGF receptors, detection, diagnosis and drug screening.

XX Sequence 448 AA;

Query Match 100.0%; Score 2407; DB 19; Length 448;

Best Local Similarity 100.0%; Pred. No. 1.le-160; Indels 0; Gaps 0;

Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OCTNGFDLDRSGQCLDIDECRTIPEACRGDMCMVNGGGLCIPRTNPVYRGPSNYS 60
 DB 26 OCTNGFDLDRSGQCLDIDECRTIPEACRGDMCMVNGGGLCIPRTNPVYRGPSNYS 85
 QY 61 TYSGSGYPAAPPLSAPNPTTISRPLCRGCGYQMDENOCVDNDECARDSHQCNPTQICI 120
 DB 86 TYSGSGYPAAPPLSAPNPTTISRPLCRGCGYQMDENOCVDNDECARDSHQCNPTQICI 145
 QY 121 NTEGGYTCSTDGYWMLLEGQCLDIDECRYGVCQQLCANVPYSCTCNPFTLNEDGRSC 180
 DB 146 NTEGGYTCSTDGYWMLLEGQCLDIDECRYGVCQQLCANVPYSCTCNPFTLNEDGRSC 205
 QY 181 QDVNCECATENPCVQTCVNTYGSFICRCDPGYTELEEDGVHSCDMECSFSEFLCOHECVNO 240
 DB 206 QDVNCECATENPCVQTCVNTYGSFICRCDPGYTELEEDGVHSCDMECSFSEFLCOHECVNO 265
 QY 241 PGTYFSCSPGTYLLDNRSCODINEGHNHHCNLOOTCNLOGGFRICDIPIRCEEYTL 300
 DB 266 PGTYFSCSPGTYLLDNRSCODINEGHNHHCNLOOTCNLOGGFRICDIPIRCEEYTL 325

QY 301 RISDNRCMCPAENPGCRDQFTLLYRDMDVYSGRSVPADIFOMQATTRYPGAYYFFQJKS 360
 DB 326 RISDNRCMCPAENPGCRDQFTLLYRDMDVYSGRSVPADIFOMQATTRYPGAYYFFQJKS 385
 QY 361 GNEGREFYMRQGPISATLVWTRPIKGPRIQLDLEMTVNTVINEFGSSVIRLRIVYSQ 420
 DB 386 GNEGREFYMRQGPISATLVWTRPIKGPRIQLDLEMTVNTVINEFGSSVIRLRIVYSQ 445
 QY 421 YPF 423
 DB 446 YPF 448

RESULT 2

AAW95709 standard; Protein; 448 AA.

AAW95709;

21-JUN-1999 (first entry)

Homo sapiens fetal kidney clone AK647 secreted protein.

Secreted protein; fetal kidney.

Homo sapiens.

WO9900405-A1.

07-JAN-1999.

29-JUN-1998; 98MO-US13530.

30-JUN-1997; 97US-0885610.

(GEMV) GENETICS INST INC.

Agostino MJ, Evans C, Jacobs K, Lavallie ER, McCoy JM,

Merberg D, Racie LA, Treacy M;

WPI; 1999-095671/08.

N-PSDB; AAX07567.

New polynucleotides encoding secreted human proteins - are derived

from foetal kidney or adult retina cDNA libraries, used as, e.g.

potential vaccines

Claim 11; Pages 52-54; 76pp; English.

The sequence is that of a secreted protein from a human fetal
 kidney clone AK296. Such a sequence is predicted to have biological
 activities which would make them suitable for treating, preventing or
 ameliorating medical conditions in humans and animals, although no
 supporting data is given. Suggested activities include nutritional
 activity, cytokine and cell proliferation/differentiation activity,
 immune stimulating (e.g. as vaccines) or suppressing activity,
 haematopoiesis regulating activity, tissue growth activity,
 activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 and thrombolytic activity, receptor/ligand activity, anti-inflammatory
 activity, cadherin/tumour invasion suppressor activity, and tumour
 inhibition activity. It is also stated to be useful for gene
 therapy.

Sequence 448 AA;

Query Match 100.0%; Score 2407; DB 20; Length 448;

Best Local Similarity 100.0%; Pred. No. 1.le-160; Indels 0; Gaps 0;

Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OCTNGFDLDRSGQCLDIDECRTIPEACRGDMCMVNGGGLCIPRTNPVYRGPSNYS 60
 DB 26 OCTNGFDLDRSGQCLDIDECRTIPEACRGDMCMVNGGGLCIPRTNPVYRGPSNYS 85

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 11:59:37 ; Search time 30.5959 Seconds
(without alignments)
1842.243 Million cell updates/sec

Title: US-09-674-379A-14
Perfect score: 2407
Sequence: 1 OCTGNFDDLRSGCCLDIDE.....INFRGSSVIRLRIVGQYF 423

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2407	100.0	448	19	AAW79739 Human EGF protein
2	2407	100.0	448	20	AAW95709 Homo sapiens fetal
3	2407	100.0	448	20	AAW94281 Human extracellular
4	2407	100.0	448	21	AAV57058 Amino acid sequenc
5	2407	100.0	448	21	AAV54989 Full length human
6	2407	100.0	448	22	AAW93573 Human polypeptide
7	2407	100.0	448	23	AAU75494 Human extracellular
8	2401	99.8	448	20	AAV08063 Human EGF-like hom
9	2401	99.8	448	22	AAU29227 Human PRO polypept
10	2401	99.8	448	22	AAW31183 Amino acid sequenc

11	2302	95.6	423	21	AAV56751
12	2302	95.6	423	21	AAV56753
13	2302	95.6	448	21	AAV56750
14	2302	95.6	448	21	AAV54990
15	2302	95.6	461	21	AAV56752
16	2302	95.6	461	21	AAV54991
17	2230	92.6	392	18	AAW31705
18	1827	75.9	335	21	AAV76008
19	1827	75.9	335	22	AAV55947
20	1827	75.9	335	22	AAV72147
21	1276.5	53.0	443	18	AAW32110
22	1276.5	53.0	443	20	AAV16587
23	1276.5	53.0	443	21	AAW33418
24	1276.5	53.0	443	21	AAV84707
25	1276.5	53.0	443	21	AAV55850
26	1276.5	53.0	443	22	AAU12330
27	1276.5	53.0	443	23	AAV84706
28	1272.5	52.9	443	21	AAV84705
29	1270.5	52.8	443	22	AAV92533
30	1237.5	51.4	433	21	AAV58353
31	1106	45.9	201	22	AAW93220
32	1093.5	45.4	493	22	AAV72892
33	1093.5	45.4	493	22	AAV84707
34	992.5	41.2	295	21	AAV84705
35	989.5	41.1	295	21	AAV84708
36	963.5	40.0	350	20	AAV08066
37	963.5	40.0	387	22	AAV68188
38	912.5	37.9	274	21	AAV76081
39	912.5	37.9	274	22	AAV56020
40	912.5	37.9	274	23	AAV72220
41	811.5	33.7	413	22	AAV80391
42	811.5	33.7	413	23	AAV65300
43	811.5	33.7	451	22	AAV80440
44	735.5	30.6	683	12	AAV11150
45	735.5	30.6	683	18	AAW27600

ALIGNMENTS

RESULT 1	AAW79739	standard; Protein; 448 AA.
ID	AAW79739	standard; Protein; 448 AA.
XX	AAW79739;	
AC	AAW79739;	
XX	25-JAN-1999	(first entry)
DT	25-JAN-1999	(first entry)
XX	Human EGF protein.	
DE	Human EGF protein.	
XX	Extracellular/epidermal growth factor-like protein; EGF; human; liver;	
KW	vascular smooth muscle cell proliferation; neurology; pathology; AIDS;	
KW	dementia; ocular; disorder; cornea; inflammation; tumor cell; kidney;	
KW	wound; epithelium; surgery; skin graft; ulcer; diabetic; trophic;	
KW	Marfan syndrome; treatment; hair follicle; embryogenesis; neoplasma;	
KW	epidermal cell; cancer; psoriasis; detection.	
XX		
OS	Homo sapiens.	
XX		
XX		
Key	Location/Qualifiers	
FT	1..25	
FT	/label= signal	
FT	26..448	
FT	/label= EGF	
FT	/note= "Extracellular/epidermal growth factor-like protein"	
FT	112..153	
FT	/label= EGF-1	
FT	154..190	
FT	/label= EGF-2	
FT	191..230	
FT	/label= EGF-3	
FT	231..271	
FT	Region	

Smooth muscle prol
Smooth muscle prol
Smooth muscle prol
Full length mouse
Smooth muscle prol
Full length mouse
Human extracellular
Rat EGF extracellu
Skin cell protein,
Rat protein isolat
Human extracellular
Extracellular prot
Human PRO226 prote
A human p53 mutat
Human S1-5 ECPM-11
Human PRO226 polyp
Human PRO226 polyp
Amino acid sequenc
Human protein sequ
Lung cancer associ
Human polypeptide,
Human EFEMP1. Hom
Human extracellular
The C-terminal of
C-terminal of p53
Human EGF-like pro
Extracellular prot
Human EGF extracel
Skin cell protein,
Human protein isol
Secreted protein e
Human albumin fusi
Gene #21 associate
Fibulin C. Homo s
Human fibulin type

TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1251 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-02251-3

Query Match 18.2%; Score 438; DB 5; Length 1251;
Best Local Similarity 29.2%; Pred No. 1.2e-26;
Matches 109; Conservative 36; Mismatches 114; Indels 114; Gaps 16;

QY 2 CTNGFDLDRSQGLDIDECRTIPAC-RGDMCVNQNGYLIPRTNPNVYRGYSNPYS 60
DB 580 CNAGYRSHPOHRYCVDVNECEA--EPCGPGKICMNTGTSYNC----- 620
QY 61 TPYSGPYPAAPPLSAPNVPITISRLICFGYQMD---ESNOCVDVDECATDSHOCNPTQ 117
DB 621 -----HCNRYRLHVAGAGRSCVDINECA-KPHLCGDG 653
QY 118 ICINTEGGYTCSTDGWYLEGQ---CLDIDECR-YGYCQO-ICANVPGSYSC-TCNPGF 171
DB 654 FCINFPGHYKNCYPGYRLKASRPICEDIDECRDPSTCPDGKCNKPGSFRCIACQPGY 713
QY 172 TLNEDGRSCODVNECATENPC-----VQT----- 195
DB 714 R-SGGGACRDNVNECSGTPCSPGWCERLPGSYRCTCAOGIRTRGRLCIDVDCEAGK 772
QY 196 -----CVNTYGSFICRCDPGYELEDGVHCSDMDECSFSEFLQHECVNQPGTYFCSCP 249
DB 773 VCODGITNTNTPGSPQCCLSGYHLSRDRSCEDIDECDFPACIGGDCINTNGSYRCLCP 832
QY 250 PGYLLDDNRSQDINECHRNHTCNLQOTCYNLQGGFKCIDPIRCEPEYLRISDNRCMC 309
DB 833 LGHRLVGGKCKDKDIDECSDPGLC-LPHACENLQGSYCV---CDEGFTLTQDQH--- 884
QY 310 PAENPGCR--DQP 320
DB 885 -----GCEVEVEP 892

Search completed: July 3, 2003, 18:26:54
Job time: 13.1125 secs

Db 674 ECLNPS-TC-PDEQVNSPGSYOCVPTGEGFRGMNQCILDVDECLPBNVCANGDCSNLEG 731
QY 162 SYSTCNPGFTLNEDGRSCODVNECATENPCVQ----- 194
Db 732 SYMCSCHKYTRTPDHRKRDIDECQGNLCVNGCKNTEGSRCTCGGQYOLSAKADQC 791
QY 195 -----TCVNTYGSFTICRCDPGYELEBDGVHCSMDSCSFSEFCQH-ECV 238
Db 792 EDIDECQHRNLCAHGSCRTNTEGSPQCVCDQGYRASLGDHCEIDINCLDKRSYVCQGRDCT 851
QY 239 NQPTGYFSCPPGYILLDNRSQDINEGHNHTCNLDQTCYNLOGGFKCT----- 290
Db 852 NTAGSYDCTCPDGF-QLDNKTCTQDINECEHPC-LCGPQGECLNTEGSEFHCVCQGGFSTIS 909
QY 291 -DPTCEEPYLRIS-----DN-----RCMC-----PAENPGCRDQPTLLYRDM 328
Db 910 ADGRGCEIDIDECVNTVCDSHGFCNTAGSFRLCTQGFQARQDQGCVD-----VNEC 963
QY 329 DVSIG 333
Db 964 ELLSG 968

RESULT 12

US-08-897-443-1
; Sequence 1, Application US/08897443
; Patent No. 5981263
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; APPLICANT: Kaaser, Matthew
; TITLE OF INVENTION: HUMAN MATRILIN-3
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/897,443
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0348 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 638 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: UTRMOT02
; CLONE: 681719
US-08-897-443-1

Query Match 18.9%; Score 455; DB 2; Length 638;
Best Local Similarity 28.1%; Pred. No. 2.7e-28;

Matches 92; Conservative 59; Mismatches 112; Indels 64; Gaps 11;
QY 1 QCTNGFDLDRSGGCLDIDECRTIPENARGDMCMYONGYLCTIRTPYRGPYSNPS 60
Db 304 QCTGYALAEGRKCAVADVICASENHGCEHE--CVNADGSLC----- 344
QY 61 TPVSGPYPAAPPLSAPNYPPIISRPLICRFGYOMD-ESNOCVDVDECATDSHOCNPTQIC 119
Db 345 -----QCHEGFALNPDKKITKTIDYCASSNHGQ--HEC 376
QY 120 INTEGGYTCGCTDGYMLLEG--CLDIDECRYG--COQLCANPGYSTCNPGETLINE 175
Db 377 VNTDSTSCQLKGFETLNPDKTCRRINYCALNPGCEHECVNMEESYCRHGYTLDP 436
QY 176 DGRSCQDVNCA--TENPCVQTCVNTYGSFTICRCDPGYELEBDGVHCSMDSCSFSEFCQ 234
Db 437 NGKTCRSHVDCAQDHDCEQCLNTEDSFVQCCEGFLINDLTCSRVYDCLLSDHCE 496
QY 235 HECVNOPTGYFSCPPGYILLDNRSQDINEGHNHTCNLDQTCYNLOGGFKCIPIR 294
Db 497 YSCVNMDSFPACCPFGHVLRSQDKTCARIDSCALGDHGC--EHSVSSSDSDFVC-----Q 550
QY 295 CEEPYLRISDNR-C---MCPAENPGC 316
Db 551 CPEGYTLREDGKTCRRKRDVCAQIDHGC 577

RESULT 13

US-08-479-722B-2
; Sequence 2, Application US/08479722B
; Patent No. 6074840
; GENERAL INFORMATION:
; APPLICANT: Bonadio, Jeffrey
; APPLICANT: Yan, Wushan
; TITLE OF INVENTION: LATENT TGF β BINDING PROTEIN (LTBP)
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Williams, Morgan & Amerson
; STREET: 7676 Hillmont, Suite 250
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77040
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,722B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US95/02251
; FILING DATE: 21-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/316,650
; FILING DATE: 30-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/199,780
; FILING DATE: 18-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fussey, Shelley P.M.
; REGISTRATION NUMBER: 39,458
; REFERENCE/DOCKET NUMBER: 4100.000500/FUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 934-7000
; TELEFAX: (713) 934-7011
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 183 amino acids
; TYPE: amino acid

NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Fastseq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/212,168
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/884,072
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0333 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 458228
US-09-212-168-5

```

1  APPLICANT: Watson, James D.
2  APPLICANT: Strachan, Lorna
3  APPLICANT: Sleeman, Matthew
4  APPLICANT: Onrust, Rene
5  APPLICANT: Murison, James Greg
6  TITLE OF INVENTION: Compositions Isolated From Skin Cells
7  TITLE OF INVENTION: and Methods For Their use
8  FILE REFERENCE: 11000.1011c1
9  CURRENT APPLICATION NUMBER: US/09/188,930A
10 CURRENT FILING DATE: 1998-11-09
11 NUMBER OF SEQ ID NOS: 348
12 SOFTWARE: FastSeq for Windows Version 3.0
13 SEQ ID NO 336
14 LENGTH: 274
15 TYPE: prt
16 ORGANISM: Human
17
18 US-09-188-930-336

```

CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,963C
FILING DATE: 11-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US96/05033
FILING DATE: 10-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF258
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-833-963C-9

Query Match 40.0%; Score 963.5; DB 2; Length 387;
Best Local Similarity 48.3%; Pred. No. 1,1e-68;
Matches 171; Conservative 55; Mismatches 123; Indels 5; Gaps 4;
QY 71 APPLSANPYPTISPLICRGYOMDESNOVDVDECATDSHOCNPTQICINTEGGYTSC 130
Db 38 ADQRIPSNP--SHRIQCAAGYEQSEHNVCODIDECTAGHNCRADQVCINLRGSAQC 95
QY 131 TDGYWLLGGCLDIDECRY-GYCOOLCANYPGYSCTCNPGFTLNEDGRSCODVNECATE 189
Db 96 PPGYQKRGECQVVIDECTIPPYCHQRCVNTPGSFYCCSGFOLAANNYTCVVIDNECDAS 155
QY 190 NPCVQTCVNTYGSFICRDPGYELEEDGVHSCMDECSFSEFLCQHECVNQPGTYFCSCP 249
Db 156 NQCAQCCYNILGSLFCQNGYELSSDRNCEDIDECRTSSYLCOYQCVNEPGRFSCMCP 215
QY 250 PGYILLDDNRSCODINECEHRNHTCNMQTCVNIQGFICIDPIRCEEPYLRISDRNCMC 309
Db 216 QGYQVY-RSRTCQDINECETNE-CREDEMCWNYHGFRCYRPNPCQDPYILLPENRCVC 273
QY 310 PAENPGCRQPPFTIYLRDMDVYSGRSVPADIFOMQATTRYPGAYVYIFQIKSGNEGREFYA 369
Db 274 PYSNMCRLPQSYIYKYSIRSRSVPSDIFQIOATTYANTINFRIRKSGNEGEFYL 333
QY 370 ROTGPISATLVWTRPIKGRREIOLDLEMTVTVINFRSSYVRLRIYSOYPF 423
Db 334 ROTSPVSAMLVYKSLSGPREHIVDEMLTVSSIGTFRISVLRLLTIIVGPSPF 387

RESULT 8
US-08-980-514-3
Sequence 3, Application US/08980514
Patent No. 6004753
GENERAL INFORMATION:
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
TITLE OF INVENTION: HUMAN SI-5-DCMP-LIKE PROT
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980,514
FILING DATE: filed herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0436 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 458228
US-08-980-514-3

Query Match 40.0%; Score 963.5; DB 3; Length 387;
Best Local Similarity 48.3%; Pred. No. 1,1e-68;
Matches 171; Conservative 55; Mismatches 123; Indels 5; Gaps 4;
QY 71 APPLSANPYPTISPLICRGYOMDESNOVDVDECATDSHOCNPTQICINTEGGYTSC 130
Db 38 ADQRIPSNP--SHRIQCAAGYEQSEHNVCODIDECTAGHNCRADQVCINLRGSAQC 95
QY 131 TDGYWLLGGCLDIDECRY-GYCOOLCANYPGYSCTCNPGFTLNEDGRSCODVNECATE 189
Db 96 PPGYQKRGECQVVIDECTIPPYCHQRCVNTPGSFYCCSGFOLAANNYTCVVIDNECDAS 155
QY 190 NPCVQTCVNTYGSFICRDPGYELEEDGVHSCMDECSFSEFLCQHECVNQPGTYFCSCP 249
Db 156 NQCAQCCYNILGSLFCQNGYELSSDRNCEDIDECRTSSYLCOYQCVNEPGRFSCMCP 215
QY 250 PGYILLDDNRSCODINECEHRNHTCNMQTCVNIQGFICIDPIRCEEPYLRISDRNCMC 309
Db 216 QGYQVY-RSRTCQDINECETNE-CREDEMCWNYHGFRCYRPNPCQDPYILLPENRCVC 273
QY 310 PAENPGCRQPPFTIYLRDMDVYSGRSVPADIFOMQATTRYPGAYVYIFQIKSGNEGREFYA 369
Db 274 PYSNMCRLPQSYIYKYSIRSRSVPSDIFQIOATTYANTINFRIRKSGNEGEFYL 333
QY 370 ROTGPISATLVWTRPIKGRREIOLDLEMTVTVINFRSSYVRLRIYSOYPF 423
Db 334 ROTSPVSAMLVYKSLSGPREHIVDEMLTVSSIGTFRISVLRLLTIIVGPSPF 387

RESULT 9
US-09-212-168-5
Sequence 5, Application US/09212168
Patent No. 6303765
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS

STREET: 9410 Key West Ave

NAME/KEY: UNSURE
 LOCATION: (141)...(141)
 NAME/KEY: UNSURE
 LOCATION: (166)...(166)
 NAME/KEY: UNSURE
 LOCATION: (167)...(167)
 NAME/KEY: UNSURE
 LOCATION: (183)...(183)
 US-09-188-930-186

Query Match 74.4%; Score 1791; DB 4; Length 337;
 Best Local Similarity 93.7%; Pred. No. 4,4e-134;
 Matches 314; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

89 RFGYMDSDNOCVDVDECATDSHOCNPQICINTEGGYCSCTDGLWLEGGCLDIDECR 148
 3 RFGYMDSDNOCVDVDECATDSHOCNPQICINTEGGYCSCTDGLWLEGGCLDIDECR 62
 149 YGYCOQLCANVPGSYSTCINPGFTLNEDGSCODVNECATENPCVQTCVNTYGSFTICRD 208
 63 YGYCOQLCANVPGSYSTCINPGFTLNEDGSCODVNECATENPCVQTCVNTYGSFTICRD 122
 209 PGYLEEDGVCNDSMDDECSFSEFLCOHECVNPGTYFCSPGYILLDNRSQDINECE 268
 123 PYELEEDGVCNDSMDDECSFSEFLCOHECVNPGTYFCSPGYILLDNRSQDINECE 182
 269 HRNHTCNLOQTCYNLOGGFKCIDPIRCEBPLYRLISDNRCMCPAENPCRDOPFTILYRDM 328
 183 XRNHTCTPLOTTCYNLOGGFKCIDPIRCEBPLYRLISDNRCMCPAENPCRDOPFTILYRDM 242
 329 DVVSGRSVPADIFOMQATTRYPGAYYIFQIKSGNEGREYMRQTGPISATLVTRPIKGP 388
 243 DVVSGRSVPADIFOMQATTRYPGAYYIFQIKSGNEGREYMRQTGPISATLVTRPIKGP 302
 389 REIOLDLEMTVNTYINFRGSSVIRLRIVYSQYPF 423
 303 RDIOLDLEMTVNTYINFRGSSVIRLRIVYSQYPF 337

DB
 OY
 US-08-833-963C-2

RESULT 4
 US-08-833-963C-2
 Sequence 2, Application US/08833963C
 Patent No. 5916769
 GENERAL INFORMATION:
 APPLICANT: Olsen, et al.
 TITLE OF INVENTION: Extracellular/Epidermal Growth Factor
 TITLE OF INVENTION: HCABA58X
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Ave
 CITY: Rockville
 STATE: MD
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/833.963C
 FILING DATE: 11-APR-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/US96/05033
 FILING DATE: 10-APR-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Brookes, A. Anders
 REGISTRATION NUMBER: 36,373
 REFERENCE/DOCKET NUMBER: PF558
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 301-309-8504

TELEFAX: 301-309-8439
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 443 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-833-963C-2

Query Match 53.0%; Score 1276.5; DB 2; Length 443;
 Best Local Similarity 52.0%; Pred. No. 2.3e-93;
 Matches 220; Conservative 67; Mismatches 119; Indels 17; Gaps 4;

1 OCTNGFPLDRSGGCLDIDECRTIPEACRGGMCMONGYLICPTPTNRYGPNYS 60
 38 ECTDGYEMDPDSQCRVYNECLTIPACKGEMKICNHYGILCLPRAAAYVINDLHG----- 93
 61 TPVSGPYAAPLAPNYPITISRPILCRFGYOMDESNOCVDVDECATDSHOCNPQICI 120
 94 ---EGP-PPVPPAQHFN-----PCPPGYEPDDQSCVDVDECAQALHDCRPSDCH 141
 121 NTBGGYCSCTDGLWLEGGCLDIDECRYGCOQLCANVPGSYSTCINPGFTLNEDGSC 180
 142 NLPGSYCTCPDGRKLGPEECVDIDECRYRQCQHRVNLPSFRCQCEPGFQLGPNNSC 201
 181 QDVNECATENPCVQTCVNTYGSFTICRDPGYLEEDGVCNDSMDDECSFSEFLCOHECVN 240
 202 VDVNECMGMAPCEBRCNNSYTFILCRHQGYELHRDSCSDIDECSSYSTLCQYRCNE 261
 241 PGTYFCSPGYILLDNRSQDINECEHRNHTCNLOQTCYNLOGGFKCIDPIRCEBPLY 300
 262 PGRFSCHCPQGYQLL-ATRLCQDIDECESGAHOCSEAOCTCVNFGYRCVPTNRCEPYI 320
 301 RISDNRCMCPAENPCRDOPFTILYRDMDVVSGRSVPADIFOMQATTRYPGAYYIFQIKS 360
 321 QVSENRCCLCPASNLCRQPSIVHRTTTSSESPADVQIDQTSVPAGAYNAFQIRA 380
 361 GNEGREYMRQTGPISATLVTRPIKGPRETOLDLEMTVNTYINFRSSVIRLRIVYSQ 420
 381 GNSGDFYIRQINNVSAMLVARVPTGPREVLDLEMTNMSYRASSVLRILVFVGA 440
 421 YPF 423
 441 YTF 443

DB
 OY
 US-08-980-514-1

RESULT 5
 US-08-980-514-1
 Sequence 1, Application US/08980514
 Patent No. 6004753
 GENERAL INFORMATION:
 APPLICANT: Yue, Henry
 APPLICANT: Guegler, Karl J.
 TITLE OF INVENTION: HUMAN SI-5-ECMP-LIKE PROT
 TITLE OF INVENTION: EIN
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Dr.
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/980.514
 FILING DATE: Filed Herewith
 PRIOR APPLICATION DATA:

Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 QCTNGFDLDRSGGCLDIDECRTIPEACRGMMCVNONGGYLCIRTNVYGGPSNPS 60
26 QCTNGFDLDRSGGCLDIDECRTIPEACRGMMCVNONGGYLCIRTNVYGGPSNPS 85
61 TPYSGPYPAAPPLSAPNYPITISRLICRFGYQMDENQCVYVDECATDSHOCNPTQICI 120
86 TPYSGPYPAAPPLSAPNYPITISRLICRFGYQMDENQCVYVDECATDSHOCNPTQICI 145
121 NTEGGYTCSTGDYGLWLEGGCLDIDECRYGYCQQLCANVPGSYSCTCNPGFTLNEDEGRSC 180
146 NTEGGYTCSTGDYGLWLEGGCLDIDECRYGYCQQLCANVPGSYSCTCNPGFTLNEDEGRSC 205
181 QDVNECATENPCVOTCVNTYGSFICRDPGYELEEDGYHCSMDSCSFSEFLCQHECVNQ 240
206 QDVNECATENPCVOTCVNTYGSFICRDPGYELEEDGYHCSMDSCSFSEFLCQHECVNQ 265
241 PGTYFCSCPGYYILLDNRSCQDINECHRNHTCMLOQTCVNLGGFKCIDPIRCEEPYL 300
266 PGTYFCSCPGYYILLDNRSCQDINECHRNHTCMLOQTCVNLGGFKCIDPIRCEEPYL 325
301 RISDNRCMCPAENPGCRDQPTIILYRDMVYSGRSVPADIFQMAATTRYPGAYYIFQIKS 360
326 RISDNRCMCPAENPGCRDQPTIILYRDMVYSGRSVPADIFQMAATTRYPGAYYIFQIKS 385
361 GNEGREPYMROTGISATLVMTPIKGPRIQLDLEMTVNTVINFRGSSVIRLRIYVSQ 420
386 GNEGREPYMROTGISATLVMTPIKGPRIQLDLEMTVNTVINFRGSSVIRLRIYVSQ 445
421 YPF 423
446 YPF 448

RESULT 2
US-09-212-168-1
; Sequence 1, Application US/09212168
; Patent No. 6303765
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/212,168
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/884,072
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0333 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
; LENGTH: 448 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: CORNNOT01
; CLONE: 45517
US-09-212-168-1

Query Match 100.0%; Score 2407; DB 4; Length 448;
Best Local Similarity 100.0%; Pred. No. 1,4e-182;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 QCTNGFDLDRSGGCLDIDECRTIPEACRGMMCVNONGGYLCIRTNVYGGPSNPS 60
26 QCTNGFDLDRSGGCLDIDECRTIPEACRGMMCVNONGGYLCIRTNVYGGPSNPS 85
61 TPYSGPYPAAPPLSAPNYPITISRLICRFGYQMDENQCVYVDECATDSHOCNPTQICI 120
86 TPYSGPYPAAPPLSAPNYPITISRLICRFGYQMDENQCVYVDECATDSHOCNPTQICI 145
121 NTEGGYTCSTGDYGLWLEGGCLDIDECRYGYCQQLCANVPGSYSCTCNPGFTLNEDEGRSC 180
146 NTEGGYTCSTGDYGLWLEGGCLDIDECRYGYCQQLCANVPGSYSCTCNPGFTLNEDEGRSC 205
181 QDVNECATENPCVOTCVNTYGSFICRDPGYELEEDGYHCSMDSCSFSEFLCQHECVNQ 240
206 QDVNECATENPCVOTCVNTYGSFICRDPGYELEEDGYHCSMDSCSFSEFLCQHECVNQ 265
241 PGTYFCSCPGYYILLDNRSCQDINECHRNHTCMLOQTCVNLGGFKCIDPIRCEEPYL 300
266 PGTYFCSCPGYYILLDNRSCQDINECHRNHTCMLOQTCVNLGGFKCIDPIRCEEPYL 325
301 RISDNRCMCPAENPGCRDQPTIILYRDMVYSGRSVPADIFQMAATTRYPGAYYIFQIKS 360
326 RISDNRCMCPAENPGCRDQPTIILYRDMVYSGRSVPADIFQMAATTRYPGAYYIFQIKS 385
361 GNEGREPYMROTGISATLVMTPIKGPRIQLDLEMTVNTVINFRGSSVIRLRIYVSQ 420
386 GNEGREPYMROTGISATLVMTPIKGPRIQLDLEMTVNTVINFRGSSVIRLRIYVSQ 445
421 YPF 423
446 YPF 448

RESULT 3
US-09-188-930-186
; Sequence 186, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Mathew
; APPLICANT: Onrust, Rene
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 186
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Rat
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (14)...(124)
; NAME/KEY: UNSURE
; LOCATION: (135)...(135)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 17:54:52 ; Search time 13.1125 Seconds
(Without alignments)
949.162 Million cell updates/sec

Title: US-09-674-379A-14

Perfect score: 2407

Sequence: 1 OCTNGFDLRQSGQCLDIDE.....INFRGSSVIRLRIVSQYF 423

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCrUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2407	100.0	448	2	US-08-884-072-1
2	2407	100.0	448	4	US-09-212-168-1
3	1791	74.4	337	4	US-09-188-930-186
4	1276.5	53.0	443	2	US-08-833-963C-2
5	1276.5	53.0	443	3	US-08-980-514-1
6	963.5	40.0	387	2	US-08-884-072-5
7	963.5	40.0	387	2	US-08-833-963C-9
8	963.5	40.0	387	3	US-08-980-514-3
9	963.5	40.0	387	4	US-09-212-168-5
10	912.5	37.9	274	4	US-09-188-930-336
11	539	22.4	1394	6	517197-30
12	455	18.9	638	2	US-08-897-443-1
13	453	18.8	1833	3	US-08-479-722B-2
14	453	18.8	1833	3	US-08-479-722B-2
15	438	18.2	1251	5	PCr-US95-02251-13
16	438	18.2	1252	1	US-08-199-780-3
17	438	18.2	1252	2	US-08-316-650-3
18	438	18.2	1253	3	US-08-479-722B-4
19	431	17.9	956	2	US-08-897-443-3
20	392	16.3	886	3	US-09-110-116-3
21	372	15.5	810	2	US-08-820-170A-34
22	372	15.5	810	3	US-09-055-699-34
23	372	15.5	810	4	US-09-273-565-34
24	372	15.5	810	4	US-09-565-538-34
25	372	15.5	810	4	US-09-661-468-34
26	369	15.3	2471	1	US-08-185-432-16
27	369	15.3	2471	1	US-08-083-590A-19

28	369	15.3	2471	3	US-08-532-384-19	Sequence 19, Appl
29	369	15.3	2471	4	US-08-899-232-1	Sequence 1, Appl
30	361.5	15.0	816	2	US-08-820-170A-37	Sequence 37, Appl
31	361.5	15.0	816	3	US-09-055-699-37	Sequence 37, Appl
32	361.5	15.0	816	4	US-09-273-565-37	Sequence 37, Appl
33	361.5	15.0	816	4	US-09-565-538-37	Sequence 37, Appl
34	361.5	15.0	816	4	US-09-661-468-37	Sequence 37, Appl
35	354.5	14.7	1964	4	US-09-467-997-1	Sequence 1, Appl
36	351.5	14.6	2703	1	US-08-185-432-19	Sequence 19, Appl
37	351.5	14.6	2703	4	US-08-899-232-4	Sequence 4, Appl
38	350.5	14.6	652	2	US-08-751-305-2	Sequence 2, Appl
39	345.5	14.4	2556	1	US-08-083-590A-20	Sequence 20, Appl
40	345.5	14.4	2556	3	US-08-532-384-20	Sequence 20, Appl
41	342.5	14.2	2556	1	US-08-185-432-17	Sequence 17, Appl
42	342.5	14.2	2556	4	US-08-899-232-2	Sequence 3, Appl
43	340.5	14.1	673	1	US-08-282-141-3	Sequence 3, Appl
44	340.5	14.1	673	1	US-08-435-434-1	Sequence 1, Appl
45	340.5	14.1	673	1	US-08-435-436-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-884-072-1
Sequence 1, Application US/08884072
Patent No. 5872234
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
TITLE OR INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/884,072
FILING DATE: Herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0333 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 448 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: CORNOT01
CLONE: 45517
US-08-884-072-1
Query Match 100.0%; Score 2407; DB 2; Length 448;
Best Local Similarity 100.0%; Pred. No. 1,4e-182;

```
QY 1 QCTNGFPLDRSGGCLDIDECRTIPEACRGDMACVNONGGYLCIPRTNPNVYRGPNPYS 60
Db 26 QCTNGFPLDRSGGCLDIDECRTIPEACRGDMACVNONGGYLCIPRTNPNVYRGPNPYS 85
QY 61 TPYSGPYPAAPPLSAPNYPITISRPLICRFYQMDENOCVDVDECATDSHOCNPTQICI 120
Db 86 TPYSGPYPAAPPLSAPNYPITISRPLICRFYQMDENOCVDVDECATDSHOCNPTQICI 145
QY 121 NTEGGYTCSTDGWLLLEGCLDIDECRYGCOQLCANVPGSYSCTCNPGFTLNEDGRSC 180
Db 146 NTEGGYTCSTDGWLLLEGCLDIDECRYGCOQLCANVPGSYSCTCNPGFTLNEDGRSC 205
QY 181 QDVNECATENPCVOTCVNTYGSFICRCDPGYELEEDGYHCSMDDECSFSEFLCQHECVNO 240
Db 206 QDVNECATENPCVOTCVNTYGSFICRCDPGYELEEDGYHCSMDDECSFSEFLCQHECVNO 265
QY 241 PGTYFCSCPFGYILLDNRSCODINECEHRNHTCNLQOTCYNLQGGFKCIDPIRCEEPYL 300
Db 266 PGTYFCSCPFGYILLDNRSCODINECEHRNHTCNLQOTCYNLQGGFKCIDPIRCEEPYL 325
QY 301 RISDNRCMCPAENBPCRDOPFTIILYRMDVYSGRSVPADIFOMOATTRYPGAYYIFQIKS 360
Db 326 RISDNRCMCPAENBPCRDOPFTIILYRMDVYSGRSVPADIFOMOATTRYPGAYYIFQIKS 385
QY 361 GNEGREFFYMRQTGPISATLVTRPIKGPRIQLDLEMIVNTVINFRGSSVIRLRITYVSO 420
Db 386 GNEGREFFYMRQTGPISATLVTRPIKGPRIQLDLEMIVNTVINFRGSSVIRLRITYVSO 445
QY 421 YPF 423
Db 446 YPF 448
```

Search completed: July 3, 2003, 18:26:16
Job time : 83.5316 secs


```

; CURRENT APPLICATION NUMBER: US/10/176,757
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 408
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-757-408

Query Match
Best Local Similarity 99.8%; Score 2401; DB 9; Length 448;
Pred. No. 2,3e-159;
Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 OCTNGFDLDRSQCLDIDECRTIPEACRGDMKCVNONGGLCTIPRNPVYRGPSNYS 60
DB 26 OCTNGFDLDRSQCLDIDECRTIPEACRGDMKCVNONGGLCTIPRNPVYRGPSNYS 85
QY 61 TPYSGPYPAAPPLSAPNPYPTISRPLICRGYQWDESNOCVDVDECATDSHCNPTQICI 120
DB 86 TPYSGPYPAAPPLSAPNPYPTISRPLICRGYQWDESNOCVDVDECATDSHCNPTQICI 145
QY 121 NTEGGYTCSTGTYWLLLEGQCLDIDECRYGCOQLCANVPGSYSCCTNPFTLNEDGRSC 180
DB 146 NTEGGYTCSTGTYWLLLEGQCLDIDECRYGCOQLCANVPGSYSCCTNPFTLNEDGRSC 205
QY 181 ODVNECATENPCVQTCVNTYGSFICRCDDPGYTELEEDGVHSDMECSFSEFLCOHECVNO 240
DB 206 ODVNECATENPCVQTCVNTYGSFICRCDDPGYTELEEDGVHSDMECSFSEFLCOHECVNO 265
QY 241 PGTYFCSPPGYILLDDNRSQCLDINECEHNNHTCNLOQTCYNLQGFRCIDPICEEPLY 300
DB 266 PGTYFCSPPGYILLDDNRSQCLDINECEHNNHTCNLOQTCYNLQGFRCIDPICEEPLY 325
QY 301 RISDNRCMCPAENPGCRDOPFTILYRDMDVSGRSVPADIFOMATTRYPGAYYIFQIKS 360
DB 326 RISDNRCMCPAENPGCRDOPFTILYRDMDVSGRSVPADIFOMATTRYPGAYYIFQIKS 385
QY 361 GNEGREFYMRQTGPISATLWTRPIKGRPREIQDLEMTYNTVYINFGSSVIRLRIVSQ 420
DB 386 GNEGREFYMRQTGPISATLWTRPIKGRPREIQDLEMTYNTVYINFGSSVIRLRIVSQ 445
QY 421 YPF 423
DB 446 YPF 448

RESULT 14
US-10-176-913-408
; Sequence 408, Application US/10176913
; Publication No. US20030022298A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C66
; CURRENT APPLICATION NUMBER: US/10/176,913
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 408
; LENGTH: 448
; TYPE: PRT
```

```

; ORGANISM: Homo Sapien
US-10-176-913-408

Query Match
Best Local Similarity 99.8%; Score 2401; DB 9; Length 448;
Pred. No. 2,3e-159;
Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 OCTNGFDLDRSQCLDIDECRTIPEACRGDMKCVNONGGLCTIPRNPVYRGPSNYS 60
DB 26 OCTNGFDLDRSQCLDIDECRTIPEACRGDMKCVNONGGLCTIPRNPVYRGPSNYS 85
QY 61 TPYSGPYPAAPPLSAPNPYPTISRPLICRGYQWDESNOCVDVDECATDSHCNPTQICI 120
DB 86 TPYSGPYPAAPPLSAPNPYPTISRPLICRGYQWDESNOCVDVDECATDSHCNPTQICI 145
QY 121 NTEGGYTCSTGTYWLLLEGQCLDIDECRYGCOQLCANVPGSYSCCTNPFTLNEDGRSC 180
DB 146 NTEGGYTCSTGTYWLLLEGQCLDIDECRYGCOQLCANVPGSYSCCTNPFTLNEDGRSC 205
QY 181 ODVNECATENPCVQTCVNTYGSFICRCDDPGYTELEEDGVHSDMECSFSEFLCOHECVNO 240
DB 206 ODVNECATENPCVQTCVNTYGSFICRCDDPGYTELEEDGVHSDMECSFSEFLCOHECVNO 265
QY 241 PGTYFCSPPGYILLDDNRSQCLDINECEHNNHTCNLOQTCYNLQGFRCIDPICEEPLY 300
DB 266 PGTYFCSPPGYILLDDNRSQCLDINECEHNNHTCNLOQTCYNLQGFRCIDPICEEPLY 325
QY 301 RISDNRCMCPAENPGCRDOPFTILYRDMDVSGRSVPADIFOMATTRYPGAYYIFQIKS 360
DB 326 RISDNRCMCPAENPGCRDOPFTILYRDMDVSGRSVPADIFOMATTRYPGAYYIFQIKS 385
QY 361 GNEGREFYMRQTGPISATLWTRPIKGRPREIQDLEMTYNTVYINFGSSVIRLRIVSQ 420
DB 386 GNEGREFYMRQTGPISATLWTRPIKGRPREIQDLEMTYNTVYINFGSSVIRLRIVSQ 445
QY 421 YPF 423
DB 446 YPF 448

RESULT 15
US-10-180-552-408
; Sequence 408, Application US/10180552
; Publication No. US20030022300A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C13
; CURRENT APPLICATION NUMBER: US/10/180,552
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 408
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-552-408

Query Match
Best Local Similarity 99.8%; Score 2401; DB 9; Length 448;
Pred. No. 2,3e-159;
Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C60
; CURRENT APPLICATION NUMBER: US/10/175,752
; PRIORITY FILING DATE: 2002-06-19
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 408
; LENGTH: 448
; TYPE: PRP
; ORGANISM: Homo Sapien
; 10-175-752-408

```

Query Match 99.8%; Score 2401; DB 9; Length 448;

Best Local Similarity 99.8%; Pred. No. 2,3e-159;

Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 OCTNGDLDROSGCCLDIDECRTIPEACRGDMCMVONGYLCIRTPNRYRGPSNPS 60
DB 26 OCTNGFDLDROSGCCLDIDECRTIPEACRGDMCMVONGYLCIRTPNRYRGPSNPS 85
QY 61 TPYSGPYPAAPPLSAPNTPTISRLICRGYOMDESNOCVDBCATDSDHOCNPTQICI 120
DB 86 TPYSGPYPAAPPLSAPNTPTISRLICRGYOMDESNOCVDBCATDSDHOCNPTQICI 145
QY 121 NTEGGYTCSTGTYWLLLEGQCLDIDECRYGYCOQLCANVPSYSCTCNPFGFTLNEDGRSC 180
DB 146 NTEGGYTCSTGTYWLLLEGQCLDIDECRYGYCOQLCANVPSYSCTCNPFGFTLNEDGRSC 205
QY 181 ODVNECATENPCVOTCVNTYGSFICRCDPGYELFEDGVCSDMDCSSEFLCQHECVNQ 240
DB 206 ODVNECATENPCVOTCVNTYGSFICRCDPGYELFEDGVCSDMDCSSEFLCQHECVNQ 265
QY 241 PGTYFSCPPGYILLDDNRSCODINECEHRNHTCNLOQTCYNLOGGFKCIDIRCEEPYL 300
DB 266 PGTYFSCPPGYILLDDNRSCODINECEHRNHTCNLOQTCYNLOGGFKCIDIRCEEPYL 325
QY 301 RISDNRCMCPAENPGCRDQPTIILYRDMDVVSGRSVPADIFOMQATTRYPGAYYIFQIKS 360
DB 326 RISDNRCMCPAENPGCRDQPTIILYRDMDVVSGRSVPADIFOMQATTRYPGAYYIFQIKS 385
QY 361 GNGREFFYMRQTGPISATLVMTPIKGPRTIOLDLEMTVNTVINFRGSSVIRLRIVSQ 420
DB 386 GNGREFFYMRQTGPISATLVMTPIKGPRTIOLDLEMTVNTVINFRGSSVIRLRIVSQ 445
QY 421 YPF 423
DB 446 YPF 448

```

RESULT 12

US-10-176-482-408

; Sequence 408, Application US/10176482

; Publication No. US20030022296A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

```

; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C70
; CURRENT APPLICATION NUMBER: US/10/176,482
; PRIORITY FILING DATE: 2002-06-20
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 408
; LENGTH: 448
; TYPE: PRP
; ORGANISM: Homo Sapien
; 10-176-482-408

```

Query Match 99.8%; Score 2401; DB 9; Length 448;

Best Local Similarity 99.8%; Pred. No. 2,3e-159;

Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 OCTNGFDLDROSGCCLDIDECRTIPEACRGDMCMVONGYLCIRTPNRYRGPSNPS 60
DB 26 OCTNGFDLDROSGCCLDIDECRTIPEACRGDMCMVONGYLCIRTPNRYRGPSNPS 85
QY 61 TPYSGPYPAAPPLSAPNTPTISRLICRGYOMDESNOCVDBCATDSDHOCNPTQICI 120
DB 86 TPYSGPYPAAPPLSAPNTPTISRLICRGYOMDESNOCVDBCATDSDHOCNPTQICI 145
QY 121 NTEGGYTCSTGTYWLLLEGQCLDIDECRYGYCOQLCANVPSYSCTCNPFGFTLNEDGRSC 180
DB 146 NTEGGYTCSTGTYWLLLEGQCLDIDECRYGYCOQLCANVPSYSCTCNPFGFTLNEDGRSC 205
QY 181 ODVNECATENPCVOTCVNTYGSFICRCDPGYELFEDGVCSDMDCSSEFLCQHECVNQ 240
DB 206 ODVNECATENPCVOTCVNTYGSFICRCDPGYELFEDGVCSDMDCSSEFLCQHECVNQ 265
QY 241 PGTYFSCPPGYILLDDNRSCODINECEHRNHTCNLOQTCYNLOGGFKCIDIRCEEPYL 300
DB 266 PGTYFSCPPGYILLDDNRSCODINECEHRNHTCNLOQTCYNLOGGFKCIDIRCEEPYL 325
QY 301 RISDNRCMCPAENPGCRDQPTIILYRDMDVVSGRSVPADIFOMQATTRYPGAYYIFQIKS 360
DB 326 RISDNRCMCPAENPGCRDQPTIILYRDMDVVSGRSVPADIFOMQATTRYPGAYYIFQIKS 385
QY 361 GNGREFFYMRQTGPISATLVMTPIKGPRTIOLDLEMTVNTVINFRGSSVIRLRIVSQ 420
DB 386 GNGREFFYMRQTGPISATLVMTPIKGPRTIOLDLEMTVNTVINFRGSSVIRLRIVSQ 445
QY 421 YPF 423
DB 446 YPF 448

```

RESULT 13

US-10-176-757-408

; Sequence 408, Application US/10176757

; Publication No. US20030022297A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C86

QY 361 GNEGREFYMRQTGPISATLVMTPIKGPRIQDLEMITVNTVINFGSSVIRLRIVSQ 420
| | | | |
DB 386 GNEGREFYMRQTGPISATLVMTPIKGPRIQDLEMITVNTVINFGSSVIRLRIVSQ 445
QY 421 YPF 423
| | |
DB 446 YPF 448

RESULT 9

US-10-173-706-408
; Sequence 408, Application US/10173706
; Publication No. US20030022293A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C7
; CURRENT APPLICATION NUMBER: US/10/173,706
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 408
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-706-408

Query Match 99.8%; Score 2401; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2.3e-159;

Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 OCTNGFDLDRSGGCLDIDECRTIPACRGDMCVNONGYLCIPRTNPPYRGFSNPYS 60
| | | | |
DB 26 OCTNGFDLDRSGGCLDIDECRTIPACRGDMCVNONGYLCIPRTNPPYRGFSNPYS 85
Y 61 TPYSGPYPAAPPLSAPNPTISRPLICRFYQMDSESNQCVNDECATDTHQCNPTQICI 120
| | | | |
DB 86 TPYSGPYPAAPPLSAPNPTISRPLICRFYQMDSESNQCVNDECATDTHQCNPTQICI 145
QY 121 NTEGGYTCSTGTYMLLEGGCLDIDECRYGCOQLANVGSYSCNPGFTLNEDGRSC 180
| | | | |
DB 146 NTEGGYTCSTGTYMLLEGGCLDIDECRYGCOQLANVGSYSCNPGFTLNEDGRSC 205
QY 181 ODVNECATENPCVQTCVNTYGSFICRCDPGYELEEDGVHCSMDDECSFSEFLCOHECVNQ 240
| | | | |
DB 206 ODVNECATENPCVQTCVNTYGSFICRCDPGYELEEDGVHCSMDDECSFSEFLCOHECVNQ 265
QY 241 PGTFECSGPGYIILLDDNRSQDINCEHRNHTCNLOQTCYNLOGGKCIDPICEPBYL 300
| | | | |
DB 266 PGTFECSGPGYIILLDDNRSQDINCEHRNHTCNLOQTCYNLOGGKCIDPICEPBYL 325
QY 301 RISDNRCMCPAENPGCRDQPFITLYRDMDVSGRSVPADIFQMATTRYGAYYIFQIKS 360
| | | | |
DB 326 RISDNRCMCPAENPGCRDQPFITLYRDMDVSGRSVPADIFQMATTRYGAYYIFQIKS 385
QY 361 GNEGREFYMRQTGPISATLVMTPIKGPRIQDLEMITVNTVINFGSSVIRLRIVSQ 420
| | | | |
DB 386 GNEGREFYMRQTGPISATLVMTPIKGPRIQDLEMITVNTVINFGSSVIRLRIVSQ 445
QY 421 YPF 423
| | |
DB 446 YPF 448

RESULT 10

US-10-175-738-408
; Sequence 408, Application US/10175738
; Publication No. US20030022294A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C45
; CURRENT APPLICATION NUMBER: US/10/175,738
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 408
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-738-408

Query Match 99.8%; Score 2401; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2.3e-159;

Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 OCTNGFDLDRSGGCLDIDECRTIPACRGDMCVNONGYLCIPRTNPPYRGFSNPYS 60
| | | | |
DB 26 OCTNGFDLDRSGGCLDIDECRTIPACRGDMCVNONGYLCIPRTNPPYRGFSNPYS 85
QY 61 TPYSGPYPAAPPLSAPNPTISRPLICRFYQMDSESNQCVNDECATDTHQCNPTQICI 120
| | | | |
DB 86 TPYSGPYPAAPPLSAPNPTISRPLICRFYQMDSESNQCVNDECATDTHQCNPTQICI 145
QY 121 NTEGGYTCSTGTYMLLEGGCLDIDECRYGCOQLANVGSYSCNPGFTLNEDGRSC 180
| | | | |
DB 146 NTEGGYTCSTGTYMLLEGGCLDIDECRYGCOQLANVGSYSCNPGFTLNEDGRSC 205
QY 181 ODVNECATENPCVQTCVNTYGSFICRCDPGYELEEDGVHCSMDDECSFSEFLCOHECVNQ 240
| | | | |
DB 206 ODVNECATENPCVQTCVNTYGSFICRCDPGYELEEDGVHCSMDDECSFSEFLCOHECVNQ 265
QY 241 PGTFECSGPGYIILLDDNRSQDINCEHRNHTCNLOQTCYNLOGGKCIDPICEPBYL 300
| | | | |
DB 266 PGTFECSGPGYIILLDDNRSQDINCEHRNHTCNLOQTCYNLOGGKCIDPICEPBYL 325
QY 301 RISDNRCMCPAENPGCRDQPFITLYRDMDVSGRSVPADIFQMATTRYGAYYIFQIKS 360
| | | | |
DB 326 RISDNRCMCPAENPGCRDQPFITLYRDMDVSGRSVPADIFQMATTRYGAYYIFQIKS 385
QY 361 GNEGREFYMRQTGPISATLVMTPIKGPRIQDLEMITVNTVINFGSSVIRLRIVSQ 420
| | | | |
DB 386 GNEGREFYMRQTGPISATLVMTPIKGPRIQDLEMITVNTVINFGSSVIRLRIVSQ 445
QY 421 YPF 423
| | |
DB 446 YPF 448

RESULT 11

US-10-175-752-408
; Sequence 408, Application US/10175752
; Publication No. US20030022295A1
; GENERAL INFORMATION:

DB 146 NTEGGYTCSTGDTGWLLEGGCLDIDECRYGCOOLCANVPSYSTCCTNPGFTLNEDGRSC 205
QY 181 QDVNECATENPCVQTCVNTGSGFICRCDPGYLEBEDVHCSMDMECSSEFLCOHECVNQ 240
DB 206 QDVNECATENPCVQTCVNTGSLICRCDPGYLEBEDVHCSMDMECSSEFLCOHECVNQ 265
QY 241 PGTYFSCPPGYILLDDNRSCODINECEHRNHTCNLOOTCYNLGGFCIDPIRCEEPYL 300
DB 266 PGTYFSCPPGYILLDDNRSCODINECEHRNHTCNLOOTCYNLGGFCIDPIRCEEPYL 325
QY 301 RISDNRCMCAENPGCRDOPFTILYRDMVYSGSVADIFOMQATTTRPGAYYIFOLKS 360
DB 326 RISDNRCMCAENPGCRDOPFTILYRDMVYSGSVADIFOMQATTTRPGAYYIFOLKS 385
QY 361 GNEGREFYRQTPISATLWTRPIKPREIQLDLEMTVTNINFRSSVIRLRITYSQ 420
DB 386 GNEGREFYRQTPISATLWTRPIKPREIQLDLEMTVTNINFRSSVIRLRITYSQ 445
QY 421 YPF 423
DB 446 YPF 448
RESULT 7
US-10-176-758-408
Sequence 408, Application US/10176758
Publication No. US20030008353A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C104
CURRENT APPLICATION NUMBER: US/10/176,758
CURRENT FILING DATE: 2002-06-21
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 408
LENGTH: 448
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-758-408
Query Match 99.8%; Score 2401; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2.3e-159;
Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 OCTNGFDLDRSGQCLDIDECRTIPEACRGDMCMVONGGYLCIPRTNPNVYRGPSNPNYS 60
DB 26 OCTNGFDLDRSGQCLDIDECRTIPEACRGDMCMVONGGYLCIPRTNPNVYRGPSNPNYS 85
QY 61 TPYSGPYPAAPLAPLADNPTISRPILICRFGYQMDENOCVADVDECATDSHOCNPTQICI 120
DB 86 TPYSGPYPAAPLAPLADNPTISRPILICRFGYQMDENOCVADVDECATDSHOCNPTQICI 145
QY 121 NTEGGYTCSTGDTGWLLEGGCLDIDECRYGCOOLCANVPSYSTCCTNPGFTLNEDGRSC 180
DB 146 NTEGGYTCSTGDTGWLLEGGCLDIDECRYGCOOLCANVPSYSTCCTNPGFTLNEDGRSC 205
QY 181 QDVNECATENPCVQTCVNTGSGFICRCDPGYLEBEDVHCSMDMECSSEFLCOHECVNQ 240
DB 206 QDVNECATENPCVQTCVNTGSGFICRCDPGYLEBEDVHCSMDMECSSEFLCOHECVNQ 265
QY 241 PGTYFSCPPGYILLDDNRSCODINECEHRNHTCNLOOTCYNLGGFCIDPIRCEEPYL 300

DB 266 PGTYFSCPPGYILLDDNRSCODINECEHRNHTCNLOOTCYNLGGFCIDPIRCEEPYL 325
QY 301 RISDNRCMCAENPGCRDOPFTILYRDMVYSGSVADIFOMQATTTRPGAYYIFOLKS 360
DB 326 RISDNRCMCAENPGCRDOPFTILYRDMVYSGSVADIFOMQATTTRPGAYYIFOLKS 385
QY 361 GNEGREFYRQTPISATLWTRPIKPREIQLDLEMTVTNINFRSSVIRLRITYSQ 420
DB 386 GNEGREFYRQTPISATLWTRPIKPREIQLDLEMTVTNINFRSSVIRLRITYSQ 445
QY 421 YPF 423
DB 446 YPF 448
RESULT 8
US-10-175-737-408
Sequence 408, Application US/10175737
Publication No. US20030013153A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C150
CURRENT APPLICATION NUMBER: US/10/175,737
CURRENT FILING DATE: 2002-06-19
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 408
LENGTH: 448
TYPE: PRT
ORGANISM: Homo Sapien
US-10-175-737-408
Query Match 99.8%; Score 2401; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2.3e-159;
Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 OCTNGFDLDRSGQCLDIDECRTIPEACRGDMCMVONGGYLCIPRTNPNVYRGPSNPNYS 60
DB 26 OCTNGFDLDRSGQCLDIDECRTIPEACRGDMCMVONGGYLCIPRTNPNVYRGPSNPNYS 85
QY 61 TPYSGPYPAAPLAPLADNPTISRPILICRFGYQMDENOCVADVDECATDSHOCNPTQICI 120
DB 86 TPYSGPYPAAPLAPLADNPTISRPILICRFGYQMDENOCVADVDECATDSHOCNPTQICI 145
QY 121 NTEGGYTCSTGDTGWLLEGGCLDIDECRYGCOOLCANVPSYSTCCTNPGFTLNEDGRSC 180
DB 146 NTEGGYTCSTGDTGWLLEGGCLDIDECRYGCOOLCANVPSYSTCCTNPGFTLNEDGRSC 205
QY 181 QDVNECATENPCVQTCVNTGSGFICRCDPGYLEBEDVHCSMDMECSSEFLCOHECVNQ 240
DB 206 QDVNECATENPCVQTCVNTGSLICRCDPGYLEBEDVHCSMDMECSSEFLCOHECVNQ 265
QY 241 PGTYFSCPPGYILLDDNRSCODINECEHRNHTCNLOOTCYNLGGFCIDPIRCEEPYL 300
DB 266 PGTYFSCPPGYILLDDNRSCODINECEHRNHTCNLOOTCYNLGGFCIDPIRCEEPYL 325
QY 301 RISDNRCMCAENPGCRDOPFTILYRDMVYSGSVADIFOMQATTTRPGAYYIFOLKS 360
DB 326 RISDNRCMCAENPGCRDOPFTILYRDMVYSGSVADIFOMQATTTRPGAYYIFOLKS 385

```
;; PRIOR APPLICATION NUMBER: 09/380137
;; PRIOR FILING DATE: 1999-08-25
;; PRIOR APPLICATION NUMBER: 09/380138
;; PRIOR FILING DATE: 1999-08-25
;; PRIOR APPLICATION NUMBER: 09/380139
;; PRIOR FILING DATE: 1999-08-25
;; PRIOR APPLICATION NUMBER: 09/403296
;; PRIOR FILING DATE: 1999-10-18
;; PRIOR APPLICATION NUMBER: 09/403297
;; PRIOR FILING DATE: 1999-10-18
;; PRIOR APPLICATION NUMBER: 09/423741
;; PRIOR FILING DATE: 1999-11-10
;; PRIOR APPLICATION NUMBER: 09/423844
;; PRIOR FILING DATE: 1999-11-12
;; PRIOR APPLICATION NUMBER: 09/522342
;; PRIOR FILING DATE: 2000-03-09
;; PRIOR APPLICATION NUMBER: 09/548815
;; PRIOR FILING DATE: 2000-04-13
;; PRIOR APPLICATION NUMBER: 09/664610
;; PRIOR FILING DATE: 2000-09-18
;; PRIOR APPLICATION NUMBER: 09/665350
;; PRIOR FILING DATE: 2000-09-18
;; PRIOR APPLICATION NUMBER: 09/709238
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 09/767609
;; PRIOR FILING DATE: 2001-01-22
;; PRIOR APPLICATION NUMBER: 09/802706
;; PRIOR FILING DATE: 2001-03-09
;; PRIOR APPLICATION NUMBER: 09/808689
;; PRIOR FILING DATE: 2001-03-14
;; PRIOR APPLICATION NUMBER: 09/866028
;; PRIOR FILING DATE: 2001-05-25
;; PRIOR APPLICATION NUMBER: 09/870574
;; PRIOR FILING DATE: 2001-05-30
;; PRIOR APPLICATION NUMBER: 09/872035
;; PRIOR FILING DATE: 2001-06-01
;; PRIOR APPLICATION NUMBER: 09/886342
;; PRIOR FILING DATE: 2001-06-19
;; PRIOR APPLICATION NUMBER: PCT/US98/14552
;; PRIOR FILING DATE: 1998-07-14
;; PRIOR APPLICATION NUMBER: PCT/US98/18824
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: PCT/US98/19093
;; PRIOR FILING DATE: 1998-09-14
;; PRIOR APPLICATION NUMBER: PCT/US98/19330
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: PCT/US98/19437
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: PCT/US98/24855
;; PRIOR FILING DATE: 1998-11-20
;; PRIOR APPLICATION NUMBER: PCT/US98/25108
;; PRIOR FILING DATE: 1998-12-01
;; PRIOR APPLICATION NUMBER: PCT/US98/25190
;; PRIOR FILING DATE: 1998-11-25
;; PRIOR APPLICATION NUMBER: PCT/US99/05028
;; PRIOR FILING DATE: 1999-03-08
;; PRIOR APPLICATION NUMBER: PCT/US99/12252
;; PRIOR FILING DATE: 1999-06-02
;; PRIOR APPLICATION NUMBER: PCT/US99/20111
;; PRIOR FILING DATE: 1999-09-01
;; PRIOR APPLICATION NUMBER: PCT/US99/20594
;; PRIOR FILING DATE: 1999-09-08
;; PRIOR APPLICATION NUMBER: PCT/US99/21090
;; PRIOR FILING DATE: 1999-09-15
;; PRIOR APPLICATION NUMBER: PCT/US99/21547

Query Match          99.8%; Score 2401; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2,3e-159;
Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 OCTNGFPLDROSGOCLDIDECRTIPEACRGDMCVNONGYLCIPRTNPFYRGPSNPS 60
Db 26 OCTNGFPLDROSGOCLDIDECRTIPEACRGDMCVNONGYLCIPRTNPFYRGPSNPS 85
```

```
QY 61 TPYSGPYPAAPPLSAPNYPITISRPILCRFGYOMDESNOQVDECATDSHQNPOTICI 120
Db 86 TPYSGPYPAAPPLSAPNYPITISRPILCRFGYOMDESNOQVDECATDSHQNPOTICI 145
QY 121 NTEGGYTCSTDGWILEGOCCLDIDECRYGYCOOLCANVPSYSCOTNPGFTLNEDGRSC 180
Db 146 NTEGGYTCSTDGWILEGOCCLDIDECRYGYCOOLCANVPSYSCOTNPGFTLNEDGRSC 205
QY 181 ODVNECATENPCYQOTCVNTGSPFCRCRDBEYELLEDVYHSDMDECSFSEFLCOHECVNQ 240
Db 206 QDVNECATENPCYQOTCVNTGSPFCRCRDBEYELLEDVYHSDMDECSFSEFLCOHECVNQ 265
QY 241 PGTYFSCPPGYILLDNRSCODINCEHNNHTCNLOOTCYNLOGGFKCIDPIRCEPPL 300
Db 266 PGTYFSCPPGYILLDNRSCODINCEHNNHTCNLOOTCYNLOGGFKCIDPIRCEPPL 325
QY 301 RISDNRCMCPAENPGCRDQPFITLYRDMDVSGRSVPADIFOMQATTRYGAYYIFQIKS 360
Db 326 RISDNRCMCPAENPGCRDQPFITLYRDMDVSGRSVPADIFOMQATTRYGAYYIFQIKS 385
QY 361 GNEGREFYMRQTPISATLWMPRIKPREIOLDLEMTVNTVINRGSSVIRLRYVSQ 420
Db 386 GNEGREFYMRQTPISATLWMPRIKPREIOLDLEMTVNTVINRGSSVIRLRYVSQ 445
QY 421 YPF 423
Db 446 YPF 448
```

RESULT 6

```
US-10-174-590-408
; Sequence 408, Application US/10174590
; Publication No. US20030008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Goddowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1CA2
; CURRENT APPLICATION NUMBER: US/10/174,590
; PRIOR FILING DATE: 2002-06-18
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 408
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-590-408
```

```
Query Match          99.8%; Score 2401; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2,3e-159;
Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 OCTNGFPLDROSGOCLDIDECRTIPEACRGDMCVNONGYLCIPRTNPFYRGPSNPS 60
Db 26 OCTNGFPLDROSGOCLDIDECRTIPEACRGDMCVNONGYLCIPRTNPFYRGPSNPS 85
QY 61 TPYSGPYPAAPPLSAPNYPITISRPILCRFGYOMDESNOQVDECATDSHQNPOTICI 120
Db 86 TPYSGPYPAAPPLSAPNYPITISRPILCRFGYOMDESNOQVDECATDSHQNPOTICI 145
QY 121 NTEGGYTCSTDGWILEGOCCLDIDECRYGYCOOLCANVPSYSCOTNPGFTLNEDGRSC 180
```

Oy 361 GNEGRFYHROTGPISATLVTRPIKGPRIQLDLEMTVTVINFRSSVIRLIYYSO 420
Db 386 GNEGRFYHMQTGPISATLVTRPIKGPRIQLDLEMTVTVINFRSSVIRLIYYSO 445

Oy 421 YPF 423
Db 446 YPF 448

RESULT 5

US-10-066-500-15
Sequence 15, Application US/10066500
Patent No. US20020177165A1

GENERAL INFORMATION:

APPLICANT: Avi J. Ashkenazi
APPLICANT: Kevin P. Baker
APPLICANT: David A. Botstein
APPLICANT: Luc Desnoyers
APPLICANT: Dan L. Eaton
APPLICANT: Napoleone Ferrara
APPLICANT: Sherman Fong
APPLICANT: Wei-Qiang Gao
APPLICANT: Hanspeter Gerber
APPLICANT: Mary E. Gerltzen
APPLICANT: Audrey Goddard
APPLICANT: Paul J. Godowski
APPLICANT: Austin L. Gurney
APPLICANT: Ivar J. Kljavin
APPLICANT: Jennie P. Mather
APPLICANT: Mary A. Napier
APPLICANT: James Pan
APPLICANT: Nicholas F. Paoni
APPLICANT: Margaret Ann Roy
APPLICANT: Timothy A. Stewart
APPLICANT: Daniel Tumas
APPLICANT: Colin K. Watanabe
APPLICANT: P. Mickey Williams
APPLICANT: William I. Wood
APPLICANT: Zemin Zang

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS ENCODING THE SAME

FILE REFERENCE: F313ORIC

CURRENT APPLICATION NUMBER: US/10/066, 500

CURRENT FILING DATE: 2002-02-01

PRIOR APPLICATION NUMBER: 10/002,796

PRIOR FILING DATE: 2001-11-15

PRIOR APPLICATION NUMBER: 60/056974

PRIOR FILING DATE: 1997-08-26

PRIOR APPLICATION NUMBER: 60/059115

PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/059263

PRIOR FILING DATE: 1997-09-18

PRIOR APPLICATION NUMBER: 60/059588

PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/062285

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/062816

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/063082

PRIOR FILING DATE: 1997-10-31

PRIOR APPLICATION NUMBER: 60/063329

PRIOR FILING DATE: 1997-10-27

PRIOR APPLICATION NUMBER: 60/063733

PRIOR FILING DATE: 1997-10-29

PRIOR APPLICATION NUMBER: 60/066364

PRIOR FILING DATE: 1997-11-21

PRIOR APPLICATION NUMBER: 60/066840

PRIOR FILING DATE: 1997-11-25

PRIOR APPLICATION NUMBER: 60/069694

PRIOR FILING DATE: 1997-12-16

PRIOR APPLICATION NUMBER: 60/074086

PRIOR FILING DATE: 1998-02-09

PRIOR APPLICATION NUMBER: 60/074092

PRIOR FILING DATE: 1998-02-09

PRIOR APPLICATION NUMBER: 60/079294

PRIOR FILING DATE: 1998-03-25

PRIOR APPLICATION NUMBER: 60/081049

PRIOR FILING DATE: 1998-04-08

PRIOR APPLICATION NUMBER: 60/095998

PRIOR FILING DATE: 1998-08-10

PRIOR APPLICATION NUMBER: 60/097000

PRIOR FILING DATE: 1998-08-18

PRIOR APPLICATION NUMBER: 60/099601

PRIOR FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: 60/099803

PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 60/099811

PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 60/099812

PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 60/100858

PRIOR FILING DATE: 1998-09-17

PRIOR APPLICATION NUMBER: 60/101922

PRIOR FILING DATE: 1998-09-24

PRIOR APPLICATION NUMBER: 60/106032

PRIOR FILING DATE: 1998-10-28

PRIOR APPLICATION NUMBER: 60/109304

PRIOR FILING DATE: 1998-11-20

PRIOR APPLICATION NUMBER: 60/125778

PRIOR FILING DATE: 1999-03-23

PRIOR APPLICATION NUMBER: 60/139695

PRIOR FILING DATE: 1999-06-15

PRIOR APPLICATION NUMBER: 60/145070

PRIOR FILING DATE: 1999-07-20

PRIOR APPLICATION NUMBER: 60/145698

PRIOR FILING DATE: 1999-07-26

PRIOR APPLICATION NUMBER: 60/149396

PRIOR FILING DATE: 1999-08-17

PRIOR APPLICATION NUMBER: 60/169495

PRIOR FILING DATE: 1999-12-07

PRIOR APPLICATION NUMBER: 08/918874

PRIOR FILING DATE: 1997-08-26

PRIOR APPLICATION NUMBER: 08/933821

PRIOR FILING DATE: 1997-09-19

PRIOR APPLICATION NUMBER: 08/960507

PRIOR FILING DATE: 1997-10-29

PRIOR APPLICATION NUMBER: 09/114844

PRIOR FILING DATE: 1998-07-14

PRIOR APPLICATION NUMBER: 09/136801

PRIOR FILING DATE: 1998-08-19

PRIOR APPLICATION NUMBER: 09/136804

PRIOR FILING DATE: 1998-08-19

PRIOR APPLICATION NUMBER: 09/136828

PRIOR FILING DATE: 1998-08-19

PRIOR APPLICATION NUMBER: 09/158342

PRIOR FILING DATE: 1998-09-21

PRIOR APPLICATION NUMBER: 09/180997

PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 09/202088

PRIOR FILING DATE: 1998-12-08

PRIOR APPLICATION NUMBER: 09/254311

PRIOR FILING DATE: 1999-03-03

PRIOR APPLICATION NUMBER: 09/254460

PRIOR FILING DATE: 1999-03-09

PRIOR APPLICATION NUMBER: 09/254465

PRIOR FILING DATE: 1999-03-05

PRIOR APPLICATION NUMBER: 09/284663

PRIOR FILING DATE: 1999-04-15

PRIOR APPLICATION NUMBER: 09/332928

PRIOR FILING DATE: 1999-06-14

PRIOR APPLICATION NUMBER: 09/332929

PRIOR FILING DATE: 1999-06-14

PRIOR APPLICATION NUMBER: 09/333075

PRIOR FILING DATE: 1999-06-14

PRIOR APPLICATION NUMBER: 09/333077

PRIOR FILING DATE: 1999-06-14

STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/275,805
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/839,525
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF224
TELEPHONE: (301) 309-8509
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 448 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-275-805-2

Query Match 100.0%; Score 2407; DB 10; Length 448;
Best Local Similarity 100.0%; Pred. No. 8.9e-160;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 OCTNGFDLDRSGGCLDIDECRTIPACRGDMVCVNONGYLCIPRTNRYRGPSNPS 60
DB OCTNGFDLDRSGGCLDIDECRTIPACRGDMVCVNONGYLCIPRTNRYRGPSNPS 85
QY TPYSGPYPAAPPLSAPNPTISRPLICRFGYOMESNOCVYDDECATDSHOCNPTQICI 120
DB TPYSGPYPAAPPLSAPNPTISRPLICRFGYOMESNOCVYDDECATDSHOCNPTQICI 145
QY 121 NTEGGYTCSDTGWMLLEGQCLDIDECRYGCOQLCANVPGSYSCNPGFTLNEDGRSC 180
DB 146 NTEGGYTCSDTGWMLLEGQCLDIDECRYGCOQLCANVPGSYSCNPGFTLNEDGRSC 205
QY 181 QDVNECATENPCVQTCVNTYGSFICRCDPGYLELEDGVHCSMDDECSFSEFLCOHECVNQ 240
DB 206 QDVNECATENPCVQTCVNTYGSFICRCDPGYLELEDGVHCSMDDECSFSEFLCOHECVNQ 265
QY 241 PGTYFCSCPBGYILDDNRSQDINECEHRNHTCNLOQTCYNLQGGFKCIDPRCEEPYL 300
DB 266 PGTYFCSCPBGYILDDNRSQDINECEHRNHTCNLOQTCYNLQGGFKCIDPRCEEPYL 325
QY 301 RISDNRCMCPAENPGCRDQPTILYRDMVYSGRSVPADIFQOMATTRYPGAYYIFQIKS 360
DB 326 RISDNRCMCPAENPGCRDQPTILYRDMVYSGRSVPADIFQOMATTRYPGAYYIFQIKS 385
QY 361 GNEGREFYMRQTPISATLVMTKPGREIQLDEMTVNTVINFRGSSVIRLRITYSQ 420
DB 386 GNEGREFYMRQTPISATLVMTKPGREIQLDEMTVNTVINFRGSSVIRLRITYSQ 445
QY 421 YPF 423
DB 446 YPF 448

RESULT 4
US-09-836-561-1
; Sequence 1, Application US/09836561
; Patent No. US20020038006A1

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
Corley, Neil C.
Guegler, Karl J.
TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/836,561
FILING DATE: 16-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/212,168
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0333 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 448 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: CORNNOT01
CLONE: 45517
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-836-561-1

Query Match 100.0%; Score 2407; DB 10; Length 448;
Best Local Similarity 100.0%; Pred. No. 8.9e-160;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 OCTNGFDLDRSGGCLDIDECRTIPACRGDMVCVNONGYLCIPRTNRYRGPSNPS 60
DB 26 OCTNGFDLDRSGGCLDIDECRTIPACRGDMVCVNONGYLCIPRTNRYRGPSNPS 85
QY 61 TPYSGPYPAAPPLSAPNPTISRPLICRFGYOMESNOCVYDDECATDSHOCNPTQICI 120
DB 86 TPYSGPYPAAPPLSAPNPTISRPLICRFGYOMESNOCVYDDECATDSHOCNPTQICI 145
QY 121 NTEGGYTCSDTGWMLLEGQCLDIDECRYGCOQLCANVPGSYSCNPGFTLNEDGRSC 180
DB 146 NTEGGYTCSDTGWMLLEGQCLDIDECRYGCOQLCANVPGSYSCNPGFTLNEDGRSC 205
QY 181 QDVNECATENPCVQTCVNTYGSFICRCDPGYLELEDGVHCSMDDECSFSEFLCOHECVNQ 240
DB 206 QDVNECATENPCVQTCVNTYGSFICRCDPGYLELEDGVHCSMDDECSFSEFLCOHECVNQ 265
QY 241 PGTYFCSCPBGYILDDNRSQDINECEHRNHTCNLOQTCYNLQGGFKCIDPRCEEPYL 300
DB 266 PGTYFCSCPBGYILDDNRSQDINECEHRNHTCNLOQTCYNLQGGFKCIDPRCEEPYL 325
QY 301 RISDNRCMCPAENPGCRDQPTILYRDMVYSGRSVPADIFQOMATTRYPGAYYIFQIKS 360
DB 326 RISDNRCMCPAENPGCRDQPTILYRDMVYSGRSVPADIFQOMATTRYPGAYYIFQIKS 385

; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 448 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: No. US20020165151a1 Relevant
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-10-041-016-2

Query Match 100.0%; Score 2407; DB 9; Length 448;
 Best Local Similarity 100.0%; Pred. No. 8.9e-160;
 Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QCTNGFPLDRSGGCLDIDECRTIPEACRGDMCMYONGGYLCIRPTNVPYRGPSNPYS 60
 |||||
 DB 26 QCTNGFPLDRSGGCLDIDECRTIPEACRGDMCMYONGGYLCIRPTNVPYRGPSNPYS 85
 QY 61 TPYSGPYPAAPPLSAPNYPPTISRPLICRFGYQMDNSQCVVDDECATDSHCNPYQICI 120
 |||||
 DB 86 TPYSGPYPAAPPLSAPNYPPTISRPLICRFGYQMDNSQCVVDDECATDSHCNPYQICI 145
 QY 121 NTEGGYTCSTGTYWLEGGCLDIDECRYGCOQLCANVPGSYCTCNGFTLNEDGRSC 180
 |||||
 DB 146 NTEGGYTCSTGTYWLEGGCLDIDECRYGCOQLCANVPGSYCTCNGFTLNEDGRSC 205
 QY 181 QDVNECATENPCVQTCVNTYGSFICRCDPGYELEDGYHCSMDSCSFSEFLCOHCVNO 240
 |||||
 DB 206 QDVNECATENPCVQTCVNTYGSFICRCDPGYELEDGYHCSMDSCSFSEFLCOHCVNO 265
 QY 241 PGTYFCSCPPIYLLDNRSCODINECHRNHTCNLQOTCYNLGGFKCIDIRCEPYL 300
 |||||
 DB 266 PGTYFCSCPPIYLLDNRSCODINECHRNHTCNLQOTCYNLGGFKCIDIRCEPYL 325
 QY 301 RISDNRCMCAENPCRDQPTIILYRDMDVYSGRSVPADIFQMOATTPRGAYYIFQIKS 360
 |||||
 DB 326 RISDNRCMCAENPCRDQPTIILYRDMDVYSGRSVPADIFQMOATTPRGAYYIFQIKS 385
 QY 361 GNEGREFYMRQTGPISATLVMTRPKGPRIQLDLEMTVNTVINFRGSSVRLRLIYVSQ 420
 |||||
 DB 386 GNEGREFYMRQTGPISATLVMTRPKGPRIQLDLEMTVNTVINFRGSSVRLRLIYVSQ 445
 QY 421 YPF 423
 |||||
 DB 446 YPF 448

RESULT 2

US-09-083-002-2
 ; Sequence 2, Application US/09083002
 ; Patent No. US20010016650A1

GENERAL INFORMATION:

; APPLICANT: Jacobs, Kenneth
 ; APPLICANT: McCoy, John M.
 ; APPLICANT: Racine, Lisa A.
 ; APPLICANT: Lavallee, Edward R.
 ; APPLICANT: Merberg, David
 ; APPLICANT: Treacy, Maurice
 ; APPLICANT: Evans, Cheryl
 ; APPLICANT: Agostino, Michael
 ; APPLICANT: Lu, Zhijian
 ; APPLICANT: Honjo, Yasuko
 ; APPLICANT: Tashiro, Kei
 ; APPLICANT: Nakamura, Tomoyuki
 ; TITLE OF INVENTION: SECRETED PROTEINS
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genetics Institute, Inc.
 ; STREET: 87 Cambridgepark Drive
 ; CITY: Cambridge
 ; STATE: MA
 ; COUNTRY: U.S.A.
 ; ZIP: 02140

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/083.002
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sprunger, Suzanne A.
 ; REGISTRATION NUMBER: P-41,323
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 498-8284
 ; TELEFAX: (617) 876-5851
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 448 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: not relevant
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein

Query Match 100.0%; Score 2407; DB 10; Length 448;
 Best Local Similarity 100.0%; Pred. No. 8.9e-160;
 Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QCTNGFPLDRSGGCLDIDECRTIPEACRGDMCMYONGGYLCIRPTNVPYRGPSNPYS 60
 |||||
 DB 26 QCTNGFPLDRSGGCLDIDECRTIPEACRGDMCMYONGGYLCIRPTNVPYRGPSNPYS 85
 QY 61 TPYSGPYPAAPPLSAPNYPPTISRPLICRFGYQMDNSQCVVDDECATDSHCNPYQICI 120
 |||||
 DB 86 TPYSGPYPAAPPLSAPNYPPTISRPLICRFGYQMDNSQCVVDDECATDSHCNPYQICI 145
 QY 121 NTEGGYTCSTGTYWLEGGCLDIDECRYGCOQLCANVPGSYCTCNGFTLNEDGRSC 180
 |||||
 DB 146 NTEGGYTCSTGTYWLEGGCLDIDECRYGCOQLCANVPGSYCTCNGFTLNEDGRSC 205
 QY 181 QDVNECATENPCVQTCVNTYGSFICRCDPGYELEDGYHCSMDSCSFSEFLCOHCVNO 240
 |||||
 DB 206 QDVNECATENPCVQTCVNTYGSFICRCDPGYELEDGYHCSMDSCSFSEFLCOHCVNO 265
 QY 241 PGTYFCSCPPIYLLDNRSCODINECHRNHTCNLQOTCYNLGGFKCIDIRCEPYL 300
 |||||
 DB 266 PGTYFCSCPPIYLLDNRSCODINECHRNHTCNLQOTCYNLGGFKCIDIRCEPYL 325
 QY 301 RISDNRCMCAENPCRDQPTIILYRDMDVYSGRSVPADIFQMOATTPRGAYYIFQIKS 360
 |||||
 DB 326 RISDNRCMCAENPCRDQPTIILYRDMDVYSGRSVPADIFQMOATTPRGAYYIFQIKS 385
 QY 361 GNEGREFYMRQTGPISATLVMTRPKGPRIQLDLEMTVNTVINFRGSSVRLRLIYVSQ 420
 |||||
 DB 386 GNEGREFYMRQTGPISATLVMTRPKGPRIQLDLEMTVNTVINFRGSSVRLRLIYVSQ 445
 QY 421 YPF 423
 |||||
 DB 446 YPF 448

RESULT 3

US-09-275-805-2
 ; Sequence 2, Application US/09275805
 ; Patent No. US20010051358A1

GENERAL INFORMATION:

; APPLICANT: OLSEN, HENRIK S.
 ; APPLICANT: LI, HAODONG
 ; TITLE OF INVENTION: EXTRACELLULAR EPIDERMAL GROWTH FACTOR
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: HUMAN GENOME SCIENCES, INC.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 17:53:42 ; Search time 83.5316 Seconds

(without alignments)
582.319 Million cell updates/sec

Title: US-09-674-379a-14

Perfect score: 2407

Sequence: 1 OCTNCFDIDRQSGCCLDIDE.....INFRGSSVIRLRIVSQYF 423

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 440863 seqs, 114992915 residues

Total number of hits satisfying chosen parameters: 440863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppa/PCUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep:*
9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*
11: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2407	100.0	448	9	US-10-041-016-2
2	2407	100.0	448	10	US-09-083-002-2
3	2407	100.0	448	10	US-09-275-805-2
4	2407	100.0	448	10	US-09-836-561-1
5	2401	99.8	448	9	US-10-066-500-15
6	2401	99.8	448	9	US-10-174-590-408
7	2401	99.8	448	9	US-10-176-758-408
8	2401	99.8	448	9	US-10-175-737-408
9	2401	99.8	448	9	US-10-173-706-408
10	2401	99.8	448	9	US-10-175-738-408
11	2401	99.8	448	9	US-10-175-752-408
12	2401	99.8	448	9	US-10-176-882-408
13	2401	99.8	448	9	US-10-176-913-408
14	2401	99.8	448	9	US-10-180-552-408
15	2401	99.8	448	9	US-10-180-557-408
16	2401	99.8	448	9	US-10-173-700-408
17	2401	99.8	448	9	US-10-174-572-408
18	2401	99.8	448	9	US-10-174-579-408
19	2401	99.8	448	9	US-10-174-579-408

20	2401	99.8	448	9	US-10-174-582-408
21	2401	99.8	448	9	US-10-174-588-408
22	2401	99.8	448	9	US-10-175-739-408
23	2401	99.8	448	9	US-10-175-740-408
24	2401	99.8	448	9	US-10-175-743-408
25	2401	99.8	448	9	US-10-176-488-408
26	2401	99.8	448	9	US-10-176-492-408
27	2401	99.8	448	9	US-10-176-747-408
28	2401	99.8	448	9	US-10-176-750-408
29	2401	99.8	448	9	US-10-176-985-408
30	2401	99.8	448	9	US-10-176-987-408
31	2401	99.8	448	9	US-10-176-991-408
32	2401	99.8	448	9	US-10-176-992-408
33	2401	99.8	448	9	US-10-176-993-408
34	2401	99.8	448	9	US-10-184-658-408
35	2401	99.8	448	9	US-10-002-796-15
36	2401	99.8	448	9	US-10-066-273-15
37	2401	99.8	448	9	US-10-066-494-15
38	2401	99.8	448	9	US-10-173-695-408
39	2401	99.8	448	9	US-10-173-697-408
40	2401	99.8	448	9	US-10-173-705-408
41	2401	99.8	448	9	US-10-174-576-408
42	2401	99.8	448	9	US-10-174-585-408
43	2401	99.8	448	9	US-10-174-586-408
44	2401	99.8	448	9	US-10-175-747-408
45	2401	99.8	448	9	US-10-176-481-408

ALIGNMENTS

RESULT 1
US-10-041-016-2
Sequence 2, Appli
Patent No. US20020165151A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
McCoy, John M.
Racie, Lisa A.
Lavallee, Edward R.
Merberg, David
Treacy, Maurice
Evans, Cheryl
Agostino, Michael
Lu, Zhijian
Honjo, Tasuku
TITLE OF INVENTION: SECRETED PROTEINS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/041,016
FILING DATE: 07-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/083,002
FILING DATE: 21-MAR-1998
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: P-41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851

096SC3 PRELIMINARY; PRT: 2673 AA.
ID 096SC3;
AC 096SC3;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Fibulin-6 (Fragment).
GN FIBL-6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MELANOMA;
RA Kostka G., Timpl R.;
RT "Partial sequence of fibulin-6 with a c-terminal region related to
RL domain II and III of the fibulin family."
DR EMBL; AJ306906; CAC37630.1;
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000875; Cetrophin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF-ca.
DR InterPro: IPR003006; IG_MHC.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00008; EGF; 5.
DR Pfam: PF00047; Ig; 17.
DR Pfam: PF00090; tsp_1; 6.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_5.
DR PROSITE: PS00268; CECROPHIN; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; UNKNOWN_3.
DR PROSITE: PS01187; EGF_CA; UNKNOWN_8.
DR PROSITE: PS50092; TSP1; 6.
FT NON_TER 1
SQ SEQUENCE 2673 AA; 291017 MW; BEAEC30B8340E272 CRC64;
Query Match 22.5%; Score 542; DB 4; Length 2673;
Best Local Similarity 27.1%; Pred. No. 6.8e-46;
Matches 142; Conservative 65; Mismatches 195; Indels 122; Gaps 17;
QY 2 CTNFGDLDROSGOCLDIDECRTIPACRGDMQVQNGGYLCIPRTNPYRGPNPST 61
DB 2170 CPKGLTIAADRTGCDIDECALGRHTCHAGDCDNTISYRCVRCGSGFRRT-SDGLSC 2228
QY 62 PYSGPYRAAPPLSPNPTTSRPLICRFGYQMDSESNOCVYDECATDSHCNPTQICIN 121
DB 2229 QDINECOESSPCRHORCFNAISFHCCEPGYOL-KGRKCMVNECR--ONVCRPDQCKN 2285
QY 122 TEGGYTC--SCTDGYWLE--GOCLDIDECRYGCO---QLCANVPGSYSCINPGFTLN 174
DB 2286 TRGYTKCIDLCPNKMTKAKENKTCIDIDECCKDTHOCRYNQCIENTRGSTRVCVCPRGYSQ 2345
QY 175 EDGRSCODVNEC-ATENPCVOTCVNTYGSFTICRCDPGYLEEDGVHCS----- 221
DB 2346 GVGPRCNDINECEQVPRKCAHQSNTPSFKICICPPQHLLGDGKSCAGLERLPNYGTQY 2405
QY 222 ----- 221
DB 2406 SSYNLARSPYRNHYPOOHYROYSHLYSYSEYNSRTSLSTRRTTIKTCPEGSEASH 2465
QY 222 ----DMDECSFEFLCOHECVNOPRTYFCSPGGYILLDDNRCODINECEHNTCNLQ 277
DB 2466 DTQVDIDECENTD-ACQHECKNTGFSYOCICPPGTQLTHNGKTCODIDECLEQNVHCGPN 2524
QY 278 QTCVNLGGFKCIDPIRCEEPYLRIS-DNRGM--CPAENPGCRDQFTILYRDMVVSGR 334
DB 2525 RMCENMGSGYCID-TPCRPVYQRPVSGFCLKNCPNDLECALSPALRYKLSLPGCI 2583
QY 335 SVPADIFOMQATTR---YPCAYTI-----FOIKSGNEGREFYMRQGPISATLYM 381
DB 2584 ATNODLIRLVAYTODGVVHPRPTTFLMVDDEQTVPFALRDEN-----LKGVVYT 2631

QY 382 TRPIKGR--EIQDLEMITVNTVINFRGSSVIRLRIYVSQYP 423
DB 2632 TRPLREAEYRMVRYRASSYANGTIEYQTPEIV--YLAVSAPY 2673

Search completed: July 3, 2003, 18:21:58
Job time : 32.0959 secs

```

OY 74 -LSAPNYPTISRPILICRFGYOMD---ESNOCV-----DVDECATDSHOC 113
DB 298 CLNTPPSPFCIRLSCGTSTAMDESETRNNCFLLIINTFNCKTFEVEDVDECNLGSHDG 357
OY 114 NPTQICINTEGYTCS----- 129
DB 358 GPLYQCRNMGYSYRCDAKCKGDELQNPMTGEYIDECVTGHNCGAGECVMTPGSRQOQ 417
OY 130 ----CTDGYWL--LEGGCLDIDECRYGCOOL-CANYPGISTCICNPGFTLNE----- 175
DB 418 KGNLCAHGEVNGATGCEGDYNEQOQCGSMECINPGTYKCKGPGYEFNDARKGED 477
OY 176 -----DGRSCOVNNEGATE--NPCVOTCV 197
DB 478 VDECINFAHVCDSLNAECINTIGSFECCKRPGFOLASDGRCEVDNETTGTIACGEOKCY 537
OY 198 NMYGSFICRCDPGYELEEDGVHCSMDCESEF--SEFLCOHECVNPGTYFCSCPPGYI 253
DB 538 NTPGSIYOCICDRFALGPDGTCKEDIDECISIMAGSGNDLCKMGCCINTKSYLQCPPEYK 597
OY 254 LDDNSCODINCEHNRHTCNQOTCYNLOGFKCIDPICEEPIYLAISDR----- 306
DB 598 IOPDGRTCVDVDECA-MGECAGSDKVCVNTLGSFRC-HSIDCPINYIHDSLNKRCNCRQP 655
OY 307 --CMCPAENPGCHDOPFTILYRDMDVYSGRSVP-----ADI---FOM 343
DB 656 SAGCLPEE--CKVPLFLTYQFISL--ARAVPISSHRPATLTKRVASPNADTEVNEEL 710
OY 344 QATRRYPGAYIYIQISGNEGREFYMRQTP--ISATLVMTRPIKPREIOLDL 395
DB 711 QKTTIVGAPVNLPAIRAN-----FLLQKGEKRSASAVVTLRDSLDGPQTVKLQL 759

```

RESULT 14

OYVS89 PRELIMINARY; PRT; 1394 AA.

AC 01-MAY-2000 (Tremblrel. 13, Created)
 AC 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE CG7526 protein (Fragment).
 GN CG7526.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_Taxid=7227;

SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

RA MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne B.D.,
 RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burtis K.C., Buzan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Feriz N.S., Gelpi W.M., Glasser K.,
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glaser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegyan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Moberly C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson S.M., Nelson K.A., Nixon K., Nussken D.R., Pacle J.M.,
 RA Palazzo D.M., Pittman K.A., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spter E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003558; AAF50538.1; -
 DR HSSP: P00736; IAPQ.
 DR InterPro: FBgn0035798; CG7526.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF-Ca.
 DR InterPro: IPR003410; Hyalin.
 DR InterPro: IPR00436; Sushi_SCR_CCP.
 DR InterPro: IPR001491; Thrombomodulin.
 DR Pfam: PF00008; EGF_11.
 DR Pfam: PF00084; sushi_2.
 DR PRINTS: PR00907; THROMBOMODULN.
 DR SMART: SM00032; CCP_2.
 DR SMART: SM00179; EGF_CA_9.
 DR SMART: SM00001; EGF_Like_5.
 DR PROSITE: PS00010; ASX_HYDROXYL_8.
 DR PROSITE: PS01166; EGF_2_9.
 DR PROSITE: PS01187; EGF_CA_10.
 DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
 KW NON_TER 1394 1394
 FT SEQUENCE 1394 AA; 152269 MW; CD29380E3162F68A CRC64;

Query Match

Best Local Similarity 32.7%; Score 547.5; DB 5; Length 1394;
 Matches 119; Conservative 40; Mismatches 128; Indels 77; Gaps 13;

```

OY 2 CTNGFD--LDROGOCIDIDECRTIPEACRGDMNCVQNGSYLCIPRTNPPYRGPSNP 58
DB 519 CAGGYELKIDGIRGCFDIDECGORTHCSDQMLCENLNGSYTL--CPGYALGLDH 576
OY 59 YSTPYSGPYPAAPPLSNPTISRPILICRFGYOMDSNOCVDVDECATDSHOCPTOI 118
DB 577 IVTSLNSSFYTDSTSSSETPS-----AHCTLDIDECSLANGNC--SHR 616
OY 119 CINTREGGYTSCDGYWMLLEGO--CLDIDEC--RYGYCOQLCANYPGISTCINPGFTLN 174
DB 617 CONEPGFGQACPLGYALISEMRITCODIDECDSNGQCSQLCLNPPGACCEGFEFLT 676
OY 175 EDGRSCDQVNECATE--NPCVOTCVNTYGSFTICRCDPGYELEEDGVHCSMDCECS--FSEPL 232
DB 677 PDGCGCADIDECSDGYDNCNSPICINLGLTHACACRGYELAKDKLSCLDVDECAGLSSG 736
OY 233 COHECVNQPGTYFCSCPPGYILLDDNSCO----- 262
DB 737 CSHECINKAATFEFCGCPGLYLLNDGRSCSPALVGCPRGTONSADGCAPIECNPGYTLGS 796
OY 263 -----DINEGHRHTCNLOOTCYNLOGGFCRCIDPICEEPLRLISDNK--CM-----CPAE 312
DB 797 DDXCVDDIDECQKONGGCS--HRCSTNTESSFC--SCPPGYELSDQKTCODIDECDD 850
OY 313 NPGC 316
DB 851 KTSC 854

```

RESULT 15

Db 458 DSLNKNQADGVSCIKVCTEDTECGNTREVLYFRAPVPSKTTISPLEVSRIVTHMG 517
QY 341 --FOQATRRPGAYIIIOIKSGNEGREFYMGOTGPISATLVTRPIKGPRTOLDLEMI 398
Db 518 VPFSDVMDLVGGHFRFRVGERNIG-----IYLVKRTISG-----IVERT 559
QY 399 TWTNVIINFRGSSVIR-----LRIVYSOYPF 423
Db 560 KVNHTKSRGTGVIILFNEHIIIVSKYF 589
RESULT 12
QY 090H16 PRELIMINARY: PRT: 554 AA.
AC 090H16:
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE D162H14.1 (Fibulin 1) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RL Lloyd D.;
RA Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z98047; CAB6295.1; -
DR HSP: P35555; IEMN.
DR InterPro: IPR000020; Anaphylatoxin.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR Pfam: PF01821; ANATO; 3.
DR Pfam: PF00008; EGF; 5.
DR SMART: SM00104; ANATO; 2.
DR SMART: SM00179; EGF_CA; 4.
DR SMART: SM00001; EGF_Like; 5.
DR PROSITE: PS01177; ANAPHYLATOXIN_1; 3.
DR PROSITE: PS01178; ANAPHYLATOXIN_2; 3.
DR PROSITE: PS00010; ASX_HYDROXYL; 4.
DR PROSITE: PS01186; EGF_2; 3.
KW EGF-like domain; Glycoprotein; Hydroxylation.
FT NON-TER
SQ SEQUENCE 554 AA; 59767 MW; FE285184599A2982 CRC64;
Query Match 24.2%; Score 581.5; DB 4; Length 554;
Best Local Similarity 36.7%; Pred. No. 9,7e-51;
Matches 114; Conservative 46; Mismatches 100; Indels 51; Gaps 12;
QY 1 OCTNGFDLDROSGCGLDIDECRTIPACRGDMCVNONGYLIPRTNPNYRGYSPY 60
Db 266 OCKSGFIQD--ALGNCIGINECLISAPRGICHTINTESGYTC----- 307
QY 61 TPYSGPYRAAPPLSPNPTTISRPLICRGYQM--DESNQCVVDYDCATDSHOCNPTQIC 119
Db 308 -----QKNVFN-----CGRGYHLNEGTRFCVGVDCAPAPAEPCGKGRHC 346
QY 120 INTEGGYSCCTDGYWL--LEGQCLDIDEC--RY--GYCQCLANVGSISCTNPGFTLN 174
Db 347 VNSPSPFCECTGYTFDGISIMKCVGNECQRPGLCGHKCENTLGSYLCSGSGFRLS 406
QY 175 EDGRSCQDVNECATENPCVQTCVNTYGSFICRCDDPEYLEE--DGVHCSMDDECSF--SEF 231
Db 407 VDRGRCCEGIEGSS--SPSQDECANVYGSQYCGRGYQLSDVDGYTCGEGIDECALPTGGH 465
QY 232 LCGHCVNQGPIYFSCSP--PGYILLDDNNSCODINECEHRNHTCNLOQTCTYNLOGGFKCI 290
Db 466 ICSYRCLNIPSSPFGSCSSPSGRILAPNGRCOIGIDECVYGINCSINFCNIGGFRCL 525
QY 291 DPICEPEYLR 301

Db 526 -AFECPEMYR 535
RESULT 13
QY 018026 PRELIMINARY: PRT: 798 AA.
AC 018026: Q20903;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE F56H11.1 protein.
GN F56H11.1
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea;
OC Rhabdilitae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilkinson J.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMuray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Ritken L., Roopra A., Saunders D., Showkhen R.,
RA Smailson N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierly-Mieg J., Thomas K., Vaudin M., Vaughan R., Waterson R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Woldman P.,
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [3]
RP SEQUENCE FROM N.A.
RA Lloyd C.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z68749; CAA92962.1; -
DR EMBL: Z68219; CAA92962.1; JOINED.
DR EMBL: Z68749; CAA92483.1; JOINED.
DR HSP: P16109; IESB.
DR InterPro: IPR000020; Anaphylatoxin.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_CA.
DR Pfam: PF01821; ANATO; 2.
DR Pfam: PF00008; EGF; 5.
DR SMART: SM00104; ANATO; 2.
DR SMART: SM00179; EGF_CA; 4.
DR SMART: SM00001; EGF_Like; 6.
DR PROSITE: PS01177; ANAPHYLATOXIN_1; UNKNOWN_1.
DR PROSITE: PS00010; ASX_HYDROXYL; 4.
DR PROSITE: PS01186; EGF_2; 5.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 798 AA; 87205 MW; 3BF1EE9ED548B89 CRC64;
Query Match 23.8%; Score 574; DB 5; Length 798;
Best Local Similarity 26.8%; Pred. No. 8,7e-50;
Matches 159; Conservative 49; Mismatches 172; Indels 214; Gaps 22;
QY 2 CTNGFDLDROSGCGLDIDECRTIPACRGDMCVNONGYTC-----IP---RTNPNY 51
Db 180 CRSGFDLAPGMACVDRNECLTROSPTQSEDCVNTIGGYTCQRRISRLVPHRRRANIG 239
QY 52 RGP---YSNPYSPPYPAAP----- 73
Db 240 NAFRRMRDDPYSR--AGEYREASQANTEFGCPMGWLFQHGHCVDIDECATLMDCLSOR 297

	RESULT	10		
042182	ID	042182	PRELIMINARY:	PRT: 681 AA.
AC	042182;			
DT	01-JAN-1998	(TREMBLrel. 05, Created)		
DT	01-JAN-1998	(TREMBLrel. 05, Last sequence update)		
DT	01-DEC-2001	(TREMBLrel. 19, last annotation update)		
DE	Fibulin-1 D.			
GN	FBLN1.			
OS	Brachydanio rerio (zebrafish) (Zebra danio).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;			
CC	Cyprinidae; Danio.			
NCBI_TextID=7955;				
[1]				
RN	SEQUENCE FROM N.A.			
RA	Zhang H.-Y., Lardelli M., Ekblom P.;			
RT	"Sequence of zebrafish fibulin-1 and its expression in developing			
RL	heart and other embryonic organs."			
Dev. Genes Evol.	0:0-0(1997).			
EMBL; AF013751; AAB80944.1; -.				
HSSP; P35555; IEMN.				
ZFIN; ZDB-GENE-990415-73; fbln1.				
InterPro: IPRO00020; Anaphylatoxin.				
InterPro: IPRO00052; Asx_hydroxyl.				
InterPro: IPRO00561; EGF-like.				
InterPro: IPRO01881; EGF_Ca.				
Pfam; PF01821; ANATO; 1.				
Pfam; PF00008; EGF; 4.				
SMART; SMO0104; ANATO; 2.				
SMART; SMO0179; EGF_CA; 5.				
SMART; SMO0001; EGF_like; 4.				
PROSITE; PS01177; ANAPHYLATOXIN_1; UNKNOWN_1.				
PROSITE; PS01178; ANAPHYLATOXIN_2; 1.				
PROSITE; PS00010; ASX_HYDROXYL; 3.				
PROSITE; PS01186; EGF_2; 3.				
PROSITE; PS01187; EGF_CA; 6.				
KW	Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.			
SEQUENCE	661 AA; 74459 MW; 175C966305A46699 CRC64;			

[illegible]

Db	644	LVGEVRYTKLAMNYTVNGVYSHR - NIINRYIYSEWF	681
RESULT 11			
09TZS1			
AC 09TZS1	PRELIMINARY:	PRT: 589 AA.	
DT 01-MAY-2000 (TREMBLrel. 13, Created)			
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE Fbulin-1D (Fragment).			
GN FBLN1.			
OS Caenorhabditis elegans.			
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdiltida; Rhabditoidea;			
OC Rhabdiltidae; Peloderinae; Caenorhabdiltis.			
OX NCBI_TaxID=6239;			
[1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=CBI488 HIM-8;			
RA MECLINE=99120531; PubMed=9923556;			
RA Barth J.L., Argaves K.M., Roark E.F., Little C.D., Argaves W.S.;			
RT "Identification of chicken and C. elegans fibulin-1 homologs and			
RT characterization of the C. elegans fibulin-1 gene.";			
RL Matrix Biol. 17:635-646(1998).			
DR EMBL; AF070477; AAC24035.1; -.			
DR HSSP; P16109; 1FSB.			
DR InterPro; IPR000152; Asx_hydroxyl.			
DR InterPro; IPR000561; EGF-like.			
DR InterPro; IPR001881; EGF_Ca.			
DR Pfam; PF00008; EGF_6.			
DR SMART; SM00179; EGF_CA_5.			
DR SMART; SM00001; EGF_like_4.			
DR PROSITE; PS00010; ASX_HYDROXYL_4.			
DR PROSITE; PS01186; EGF_2_5.			
DR PROSITE; PS01187; EGF_CA_7.			
FT Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.			
FT NON_TER			
SEQUENCE 589 AA; 63984 MW; 8EA3E8FCE0B97B6 CRC64;			

Query Match	24.6%	Score 592.5	DB 5	Length 589
Best Local Similarity	27.4%	Pred. No. 7.8e-52		
Matches 156	Conservative 52	Mismatches 155	Indels 207	Gaps 20
QY	2	CTNGFDDLDROSGOCLD-IDECRTIPEACRGDMCMYONGYLCIPIRTPNPNYGRPYSNPYX	60	
Db	79	CRSGFDLAPDGMACVVDHIDECATLMDDCLESGORCLNTPGSFKCI-----	122	
QY	61	TPYSGPYPAAPRLSAPNYPPLISRPILICRFGYQMD-ESNGCYVDVECATSDHQCNPTQIC	119	
Db	123	-----PRLSGTGYAMDSFTEKRCDDVDECNLGSIDCGPLVYC	159	
QY	120	INTEGGYTC-----SCTDGYWTLLEGOCCLDIDECRYG-----	151	
Db	160	RNTGGSYRCDAKKCKGSDGLQNPMTGEGCHSYTCPNKYFRKNGNCNDIDECYVGHNGAGABE	219	
QY	152	-----COQ-----LCA-----NVPGSYSTCNBP	170	
Db	220	CVNTPGSGFRCCQKGNLCAHGYEVNGATGFCEDVNECQGVGSMCINLPGYKRCKGPG	279	
QY	171	FTLNE-----DGRSCODVNEC	186	
Db	280	YEFNDAKKRCEDVDECIKFAGHVODLSAECINTIGSFRCCKKPGFQLASDGRRCRDVNEC	339	
QY	187	ATE-NPCVQTCVNTYGSFICRDPGYELEEDGVHCSMDDECSF-----SEFLQHECVNOP	241	
Db	340	TTGTAACEQCKVCNIPGSIYOCICDGFALPGDGTCEKDIDECSSIMAGSGNDLGMGCINMTK	399	
QY	242	GYTSCSPGYYILLDDNRSCODINECEHRNNTCNLQOVCYNLGGSEFKIDIPRICEPYLR	301	
Db	400	GSYLCQCPGKGIKIDPDGRTCVADVECA-MGECAGSDKVCVNTILGSFK-HSIDCEPTNYIH	457	
QY	302	ISDNR-----C-MCPAENPGC-RDQPTILYRDMDVYGRSV-PADI-----	340	

Db 554 S-LECPENYRKSGDYLKLEKTDITRCIKSCRPNDVNCVDPVHTISHTVLSPTFERETR 612
 QY 337 PADIFQMOA-TTRYPC--AYTFOIKSGNFGREF-----YKQTGPISATLWMPRIKCP 388
 Db 613 PEEILFLRAITTPYPANQADIIEDITEGNRESFDIKRYR--DGMTVGVVQVRIPIVP 670
 QY 389 REIOLDLE-ITVNTVINFRSSVIRLRIYVSQYF 423
 Db 671 FHAILKLENNYVGVVSHR--NIVNHTVSEIWF 704

RESULT 8

QY3V7 PRELIMINARY; PRT; 576 AA.

ID 09Y3V7
 AC 09Y3V7
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Hypothetical 63.3 kDa protein (Fragment).
 NM DKFZP586A1519.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=UTERUS;
 RA Wambutt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL050095; CAB43267.1; -
 DR HSSP: P00736; IAPQ.
 DR Interpro: IPR000152; Asx_hydroxyl.
 DR Interpro: IPR000561; EGF-like.
 DR Interpro: IPR001881; EGF_Ca.
 DR Pfam: PF00008; EGF_6.
 DR SMART: SM00179; EGF_CA; 8.
 DR SMART: SM00001; EGF_Like; 2.
 DR PROSITE: PS00010; ASX_HYDROXYL; 4.
 DR PROSITE: PS01186; EGF_2; 4.
 DR PROSITE: PS01187; EGF_CA; 9.
 DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation;
 KW Hypothetical protein; Repeat.
 KM NON_TER
 FT SEQUENCE 576 AA; 63274 MW; FF2F0E9B185D8AC3 CRC64;

Query Match 29.2%; Score 703.5; DB 4; Length 576;
 Best Local Similarity 34.5%; Pred. No. 3.5e-63;
 Matches 142; Conservative 58; Mismatches 152; Indels 59; Gaps 11;

1 OCTNGFDLDRSGGCLDIDECRTIPACRGDMVCVNONGYLCLIPRTNPPYRGPSNPYS 60
 Db 187 RCMDSF-LQDPKNCVDINECTSLSPCARPFSCTINTVSGYT----- 228
 QY 61 TPYSGPYAAAPPLSAPNPTISRLPICRFYQ-MDESNOCVNDVCATDSHOCNPYQIC 119
 Db 229 -----QRNPLICARGYHASDGTGKCVNDVNEGETVHNGEGEQVC 267
 QY 120 INTEGTCSCSTDGWM--LLEGQCLDIDECRG---YCOOLCANVGSYCTCPGFTLN 174
 Db 268 HNLPSYSCDCAGQFQRAFGRCIDVNCWASPGRLCOHTCENTLGSYSCASGFLA 327
 QY 175 EDGRSCQDVNECATENPCVQTCVNTYGSFICRCDDPYELEEDGVHSCDMECS-FSEFLC 233
 Db 328 ADGRKCEVNECEAOR-CSQECANITYGTCYCRGQYQIAEDGHTCTDIDECAGAGITC 386
 QY 234 QHECVNQGTGFCSCP-PGYILLDNRSQDINECHRNHTCNMQTCYNIQGGFKCIDP 292
 Db 387 TFRCLNVGSGYQACPEQGYWTANGRSCKVDYDECALGTHNCESEATCHNIQGSFRL-R 445
 QY 293 IRCEEPYLRISDNRCMCPAENPCRD-----QPTILYRDMVDVSGRSVPADIFQMAT 346
 Db 446 FCEPNNYQVSKTCC---ERTTCHDFLECONSFAITHYQINLFOGLVPAHIFRIGPA 501

QY 347 TRPGAYYIFQIKSGNEGRFYKRGTPISATLWMPRIKPREIOLDLE 397
 Db 502 PAFGTGTLALINTIKGNEGYFTGRUNAYGVVYLQRAVLEPRDFALDVEW 552

RESULT 9

QYHBO5 PRELIMINARY; PRT; 495 AA.

ID 09HBO5
 AC 09HBO5
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical 54.3 kDa protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gu J.R., Wan D.F., Zhao X.T., Zhou X.M., Jiang H.Q., Zhang P.P.,
 RA Qiu W.X., Huang Y., Qiu X.K., Qian L.F., He L.P., Li H.N., Yu Y.,
 RA Yu J., Han L.H.;
 RT "Novel Human cDNA clones with function of inhibiting cancer cell
 RT growth.";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF217999; AAG17241.1; -
 DR HSSP: P35555; IENM.
 DR Interpro: IPR000152; Asx_hydroxyl.
 DR Interpro: IPR000561; EGF-like.
 DR Interpro: IPR001881; EGF_Ca.
 DR Pfam: PF00008; EGF_6.
 DR SMART: SM00181; EGF_9.
 DR SMART: SM00179; EGF_CA; 9.
 DR SMART: SM00001; EGF_Like; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; 3.
 DR PROSITE: PS01186; EGF_2; 3.
 DR PROSITE: PS01187; EGF_CA; 8.
 KW EGF-like domain; Glycoprotein; Hydroxylation; Hypothetical protein.
 SQ SEQUENCE 495 AA; 54340 MW; C40434B6C82E3370 CRC64;

Query Match 28.0%; Score 674.5; DB 4; Length 495;
 Best Local Similarity 36.2%; Pred. No. 2.7e-60;
 Matches 137; Conservative 58; Mismatches 126; Indels 57; Gaps 15;

1 OCTNGFDLDRSGGCLDIDECRTIPACRGDMVCVNONGYLCLIPRTNPPYRGPSNPYS 60
 Db 163 QCKSGFIOD-ALGNCIDINECTSLISAPCEIGHTCINTWEGSYTC----- 204
 QY 61 TPYSGPYAAAPPLSAPNPTISRLPICRFYQ-MDESNOCVNDVCATDSHOCNPYQIC 119
 Db 205 -----QKNVFN-----CGRGYHLENGTRCVNDVDECAPPAEPCGKGHRC 243
 QY 120 INTEGTCSCSTDGYWL--LLEGQCLDIDEC-RY--GYCOOLCANVGSYCTCPGFTLN 174
 Db 244 VNSPSFCECKTGYTFDSDISMVCVNDVNCQRPGLCHGKCGNTLGSYSCSVGRRLS 303
 QY 175 EDGRSCQDVNECATENPCVQTCVNTYGSFICRCDDPYELEE-DGVHSCDMECSF--SEF 231
 Db 304 VDRSCEDINECS--SPSOECANVGSYQCYCRRGYQISLDVGVCEIDECALPTGGH 362
 QY 232 LCOHCEVNQGTGFCSCP-PGYILLDNRSQDINECHRNHTCNMQTCYNIQGGFKCI 290
 Db 363 ICSTRCLINPGSFQSCSPSSGRILAPNGNCDIDECVTGHNCSINTECFNIQGGFRL 422
 QY 291 DPIRCEEPYLRISDNRC-MCPA-ENPGCRDQPTILYRDMVDVSGRSVPADIFQMAT 348
 Db 423 -AFEPENYRNSAARICERLPCHENRECKLPLRTIYHLSPTINQARAIVYRMRPSSA 481
 QY 349 YPGAYYIFQIKSGNEGR 366
 Db 482 VPGD---SMQAGHHRQ 495

QY	61	TFYSGYGAAPAAAPPLSAPNNPTISRPLICRFGYOMESNCCVYVDECATSHQCNPTQIC	120
Db	94	---EGP-PPVPVPAQHNP-----PCPPGTEPPDDQDSCVDVDECAQALHDPSPQDCH	141
QY	121	NTGGATGCTGCTGYYMLLEGQCLDIDECRCGYCQQLCANYPGSGYCTCNGFTLNEDGASC	180
Db	142	NLSGSGYQCTCPGQYRIKIGPECYVDIDBCRRYCOHRKRVNLPGRRCQCEGFGQGPNNRSC	201
QY	181	ODVNECATENPCVQTCVNTYGSFICRCDPGYELEDGVHCSMDCECSFSEFLCQHECVNQ	240
Db	202	VDVNECDMAPEORCFNSGYTFGLCRCHQGYELHRDGFSCSDIDECSSSYLCQYRCVNE	261
QY	241	PCTYSCSPGYILLDDNRSQCODINCEHRNHTCNLOQRCYVNLGGCFKIDPTIRCEPVL	300
Db	262	PERFCHCPCQIQL-ATRLCDDIDCESGAHOCSEAQICVNFHGGRKCVDRNRCVEPYI	320
QY	301	RISDNRCMCAPENPGCRDPFTILYRDMDVGSRSVPADIFQMAATRTPCAYIIRQIS	360
Db	321	GVSENRKCLCPASNPLCREQPSIVHRVMTITSEBRMPADVPFQQAIVSPGAVNARQIA	380
QY	361	CNEGAEFFYKQTCIPISATIVMTRPIKCPREIQDLEMTIVNTYVINEGSSVRLRIYVSO	420
Db	381	GNQSDPYRTIINNVSAMLVLARPVYGPREVYLDLEMTVNTSLSYRASSVRLITFVGA	440
QY	421	YPF 423	
Db	441	YTF 443	
RESULT 4			
ID	0922K8	PRELIMINARY:	PRT: 685 AA.
AC	0922K8:		
DT	01-DEC-2001 (TREMBLrel, 19, Created)		
DT	01-DEC-2001 (TREMBLrel, 19, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel, 21, Last annotation update)		
DE	Similar to fibulin 1.		
GN	FBLN1.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RM	[1]		
RA	SEQUENCE FROM N.A.		
RP	Strausberg R.		
RL	Submitted Apr-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL, BC007140; Fbln07140.1; -.		
DR	MGD; MGI:95487; Fbln1.		
DR	InterPro: IPR000020; Anaphylatoxin.		
DR	InterPro: IPR000152; Asx_hydroxyl.		
DR	InterPro: IPR000561; EGF-like.		
DR	InterPro: IPR001861; EGF_Ca.		
DR	Pfam: PF01821; ANATO; 3.		
DR	Pfam: PF00008; EGF; 6.		
DR	PROSITE: PS01177; ANAPHYLATOXIN_1; UNKNOWN_3.		
DR	PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_4.		
DR	PROSITE: PS01186; EGF_2; UNKNOWN_3.		
DR	PROSITE: PS01187; EGF_Ca; UNKNOWN_8.		
SO	SEQUENCE 685 AA; 75283 MW; EFDD77D7F66B73B8 CRC64;		
Query Match 30.8%; Score 741.5; DB 11; Length 685;			
Best Local Similarity 36.5%; Pred. No. 5.7e-67;			
Matches 158; Conservative 64; Mismatches 152; Indels 59; Gaps 16			
QY	1	QCTNGFDLDROSGQCLDIDECRTIEPACRGDMCYNONGYLCTIPRTNPFYBGPYSNPS	60
Db	295	QCKSGFIQD-ALGNIDINIECLISAPCVGQTCINTGESSYC-----	336
QY	61	TPYSGPYAAPPLSAPNNPTISRPLICRFGYQM-DESNQCVYVDECATSHQCNPTQIC	119
Db	337	---QKNVPN-----GGRGYHNEGTRCYVDVDECSPPAPCGKNGHC	375

OY	120	INTBEGYTCSCHDGYWL--LEGOCLDIDBC-RY--GTCQQLCANPAGSISCTCNBEFTLN	174
Db	376	INSPSPFRCEKACGAYFPGDISRTCDVINCCOAPRGRLCHKCKENPGEHSCSASAFRLS	435
OY	175	EDGRSCDOVNECATENPCVQTGVNVTGSGPICRCDPRXYELE-E-DGVHCSMDDECSF--SEF	231
Db	436	VDRGRCEDEVNEC-LNSPFGSQEACANYTGSIGCYCRRGYQLSDVDGYTCEDIDEICALPTGGH	494
OY	232	LCQHBCVNOPGTGYFCSCP-PGYILLDDNESCODINECEHRNHPTCNLQOTCYNLQGCFKI	290
Db	495	ICSYRCINIPGSGFQCSPPSSGRYLRPNGRNCODIDECYGIHNCSINETCFNIQGSFRCL	554
OY	291	DPICEEPFLIRISDNRC-MCPA-ENRPGCDOFFITLYRMDDVVGSGSVADIFOMQATRR	348
Db	555	S-FECPENTRRARDRCERLPCHENOECRPLRLITYHALSEPPTNQVAVVVRMGPSSA	613
OY	349	YPGAAVIYIOLIKSGNGRGREFYMROQTGPSIALTVMTBPRIKPREIQDLDM--ITVMNVNF	406
Db	614	VPGDSMOALATITGNNEGFEFTTKVKYSHHSVALTLKPPIPERROLLTYKKMDLIRHGIVSSF	673
OY	407	RGSSVIRLRIVYS 419	
Db	674	---VAKLFIEVS 682	
 RESULT 5 Q8TBH8 PRELIMINARY; PRT; 683 AA.			
ID	Q8TBH8;		
AC	Q8TBH8;		
DT	01-JUN-2002 (TREMBLrel. 21, Created)		
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Fibulin 1.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxId=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE-BRAIN;		
RA	Strausberg R.;		
RL	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC022497; AAH22497.1; --		
SO	SEQUENCE 683 AA; 74423 MW; 2665A3961B6403B4 CRC64;		
 Query Match 30.6%; Score 735.5, DB 4; Length 683; Best Local Similarity 35.7%; Pred. No. 2,3e-66; Matches 155; Conservative 66; Mismatches 152; Indels 61; Gaps 16			
OY	1	OCTNFFLDROSGOGLDIDECTRTPEACGDMGVNONGGYLICPRTNPVYGPSNPYS 60	
Db	293	OCKSFIOD-ALGNCIDINECLISAPRCPIGHTCINTBESSYC----- 334	
OY	61	TPYSGPYPAARPLSLAPNYPTISRPLICRFGYQM-DESNQCVDVDECATDSHCNPTQIC 119	
Db	335	-----QKNVPN-----CGRGYNLBDEGRVCYDVDECAPRAPRCGKGHRG 373	
OY	120	INTBEGYTCSCHDGYWL--LEBOCLDIDIBC-RY--GYCOQLCANPAGSISCTCNBEFTLN 174	
Db	374	VNSPSPFRCEKCTGYTFPDGISRMCDVNECOQRYPRRLGCHKCENTLGSTLCSVGFRLS 433	
OY	175	EDGRSCDOVNECATENPCVQNTGVNVTGSGPICRCDGYELE-E-DGVHCSMDDECSF--SEF 231	
Db	434	VDRGRCEDEVNECS-SPOEQECANYGSIQCTCRRGYQLSDVDGYTCEDIDEICALPTGGH 492	
OY	232	LCQHBCVNOPGTGYFCSCP-PGYILLDDNRSCODINECEHRNHPTCNLQOTCYNLQGCFKI 290	
Db	493	ICSYRCINIPGSGFQCSPPSSGRYLRPNNSNOCDIDECYGIHNCSINETCFNIQGSFRCL 552	
OY	291	DPICEEPFLIRISDNRC-MCPA-ENRPGCDOFFITLYLRMDVVGSGSVADIFOMQATTR 348	
Db	553	AFCPEPVYRSAAARCBRLPCHERNRECSKLPLRTYYHLSPPTINOAPAVFERMGPPSSA 611	

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: July 3, 2003, 15:34:32 ; Search time 30.5959 seconds
(without alignments)
2848.683 Million cell updates/sec

Title: US-09-674-379A-14

Perfect score: 2407

Sequence: 1 OCTNGFDDRDGSGCCLDIDE.....INFRGSSVIRLRIVYSGYRF 423

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

otal number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp Vertebrate: *
14: sp Unclassified: *
15: sp_virus: *
16: sp_bacteriopl: *
17: sp_archaeop: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1276.5	53.0	443	4	Q96TF5
2	1273.5	52.9	443	11	Q9JMK6
3	1253.5	52.1	443	4	Q9H3D5
4	741.5	30.8	683	11	Q922K8
5	735.5	30.6	683	4	Q8TBH8
6	731.5	30.4	1174	11	Q99K58
7	713	29.6	704	13	Q73774
8	703.5	29.2	576	4	Q9Y3V7
9	674.5	28.0	495	4	Q9HBQ5
10	629.5	26.2	681	13	Q42182
11	592.5	24.6	589	5	Q9T2S1
12	581.5	24.2	554	4	Q9UH16
13	574	23.8	798	5	Q18026
14	547.5	22.7	1394	5	Q9V8S9
15	542	22.5	2673	4	Q96SC3
16	539	22.4	1394	4	Q8TD95

17	531	22.1	5636	4	Q96RW7	Q96RW7 homo sapien
18	518.5	21.5	2872	11	Q9WUH8	Q9WUH8 rattus norv
19	517.5	21.5	741	4	Q96K89	Q96K89 homo sapien
20	512.5	21.3	3857	11	Q88840	Q88840 mus musculu
21	512	21.3	1713	11	Q88349	Q88349 mus musculu
22	511.5	21.3	2809	4	Q960P8	Q960P8 homo sapien
23	508	21.1	188	11	Q8RIU8	Q8RIU8 mus musculu
24	507.5	21.1	708	13	P87363	P87363 gallus gall
25	506.5	21.0	2906	11	Q9WUH9	Q9WUH9 rattus norv
26	491.5	20.4	937	5	Q9BR11	Q9BR11 clone intes
27	484.5	20.1	1963	6	Q28019	Q28019 bos taurus
28	483.5	20.1	447	5	Q9V4B8	Q9V4B8 dirosophila
29	482.5	20.0	1764	11	Q35806	Q35806 rattus norv
30	480.5	20.0	517	4	Q9NP01	Q9NP01 homo sapien
31	474	19.7	746	4	Q96HB9	Q96HB9 homo sapien
32	474	19.7	1256	4	Q9NS15	Q9NS15 homo sapien
33	474	19.7	1382	4	Q9H7K2	Q9H7K2 homo sapien
34	471.5	19.6	1095	11	Q60784	Q60784 mus musculu
35	469.5	19.5	1511	4	Q75412	Q75412 homo sapien
36	469.5	19.5	1664	5	Q9TVQ2	Q9TVQ2 caenorhabdi
37	469.5	19.5	1821	4	Q14767	Q14767 homo sapien
38	468.5	19.5	669	4	Q75441	Q75441 homo sapien
39	468.5	19.5	1062	11	Q60789	Q60789 mus musculu
40	467	19.4	1289	5	Q8SS53	Q8SS53 dictyosteli
41	466	19.4	1587	4	Q00508	Q00508 homo sapien
42	465.5	19.3	2189	5	Q9BI05	Q9BI05 elmeria ten
43	462.5	19.2	733	5	Q25678	Q25678 podocoryne
44	460.5	19.1	961	11	Q9BQC6	Q9BQC6 mus musculu
45	455	18.9	937	4	Q96TF5	Q96TF5 homo sapien

ALIGNMENTS

RESULT 1
ID Q96TF5 PRELIMINARY; PRT; 443 AA.
AC Q96TF5;
DT 01-DEC-2001 (TREMURel. 19, Created)
DT 01-DEC-2001 (TREMURel. 19, Last sequence update)
DT 01-JUN-2002 (TREMURel. 21, Last annotation update)
DE Mutant p53 binding protein 1 (MBP1).
GN MBP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tanaka S.;
RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RA Tanaka S.; Sugimachi K.; Sugimachi K.;
RT "Human mutant p53 binding protein (MBP1).";
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AB030655; BAA2880.1; -
DR InterPro: IPR000152; ASX_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF-Ca.
DR Pfam: PF00008; EGF_4.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_4.
DR PROSITE: PS01186; EGF_2; UNKNOWN_4.
DR PROSITE: PS01187; EGF_CA; UNKNOWN_6.
SQ SEQUENCE 443 AA; 45421 MW; 9CE175F4F388A56D CRC64;
Query Match 53.0%; Score 1276.5; DB 4; Length 443;
Best Local Similarity 52.0%; Pred. No. 8e-122;
Matches 220; Conservative 67; Mismatches 119; Indels 17; Gaps 4;
QY 1 OCTNGFDDRDGSGCCLDIDE CRTIP EACRGDMCYNONGVYLCIRTPNVPYSGPSNPYS 60
DB 38 ECTDGYEMDPDSQHC RDVNECLITPEACKGEMKCIINHGYGLCPRSAVINDLHG----- 93

```

FT DISULFID 640 654 BY SIMILARITY.
FT DISULFID 657 669 BY SIMILARITY.
FT DISULFID 675 686 BY SIMILARITY.
FT DISULFID 681 695 BY SIMILARITY.
FT DISULFID 698 709 BY SIMILARITY.
FT DISULFID 715 726 BY SIMILARITY.
FT DISULFID 721 735 BY SIMILARITY.
FT DISULFID 737 750 BY SIMILARITY.
FT DISULFID 756 767 BY SIMILARITY.
FT DISULFID 762 776 BY SIMILARITY.
FT DISULFID 778 791 BY SIMILARITY.
FT DISULFID 797 808 BY SIMILARITY.
FT DISULFID 803 817 BY SIMILARITY.
FT DISULFID 819 832 BY SIMILARITY.
FT DISULFID 838 850 BY SIMILARITY.
FT DISULFID 845 859 BY SIMILARITY.
FT DISULFID 861 873 BY SIMILARITY.
FT DISULFID 879 891 BY SIMILARITY.
FT DISULFID 885 900 BY SIMILARITY.
FT DISULFID 902 915 BY SIMILARITY.
FT DISULFID 921 933 BY SIMILARITY.
FT DISULFID 927 942 BY SIMILARITY.
FT DISULFID 944 957 BY SIMILARITY.
FT DISULFID 963 975 BY SIMILARITY.
FT DISULFID 970 984 BY SIMILARITY.
FT DISULFID 986 1000 BY SIMILARITY.
FT DISULFID 1101 1114 BY SIMILARITY.
FT DISULFID 1109 1123 BY SIMILARITY.
FT DISULFID 1125 1138 BY SIMILARITY.
FT DISULFID 1144 1155 BY SIMILARITY.
FT DISULFID 1150 1164 BY SIMILARITY.
FT DISULFID 1166 1179 BY SIMILARITY.
FT DISULFID 1298 1309 BY SIMILARITY.
FT DISULFID 1304 1318 BY SIMILARITY.
FT DISULFID 1320 1333 BY SIMILARITY.
FT DISULFID 1339 1354 BY SIMILARITY.
FT DISULFID 1349 1363 BY SIMILARITY.
FT DISULFID 1365 1378 BY SIMILARITY.
FT MOD_RES 647 647 HYDROXYLATION.
FT MOD_RES 810 810 HYDROXYLATION.
FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 870 870 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 923 923 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1039 1039 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 1394 AA: 152791 MW: DFCAB8IA40B2C7D1 CRC64:
/FTId-CAR_000184.

```

Query Match 22.48; Score 539; DB 1; Length 1394;
 Best Local Similarity 29.68; Pred. No. 1e-33;
 Matches 126; Conservative 51; Mismatches 126; Indels 122; Gaps 17;

```

QY 2 CTNGFDLDRSGOGLDIDECRTIPEACRGMCMVNONGYLCIPRTNPVYRGPYSPYST 61
DB 573 CYEGYRFSEQRCKVDIDECTQVQHLG-SQGRCENTEGSFICI----- 614
QY 62 PYSGRPAAPRLSAPRYPTISRLI-----CRFGYQMDENOCVDVD 104
DB 615 -CPAGFMASEGTNCIDVDECLRPDVCGECHCVNTGAFRCGYCDSDGYRMTORGCECDID 673
QY 105 ECATDHSOCNPTQICINTEGGYTC-SCTDGYWLLGGOGLDIDEC-RYGYCOQ-ICANYPG 161
DB 674 ECLNPS-TC-PDROCVNPSGTYQVCTEGFRGNNGOGLDIDVDECLFPVNCANGDCSNLEG 731
QY 162 SYSCGCPGFTLNEDGRSCQDVNCAENPCVQ----- 194
DB 732 SYMCSCHKGYTRPDHCHRDIDECQOGLCNVNGCKNTGSRFCTCGGYQLSAKDDC 791
QY 195 -----TCVNTYGSFICRCDREYLEBDGVHCSMDDECSFSEFLQH-ECV 238
DB 792 EDIDECQHRHLCAHGOCRTNTEGSPQCVDCGYRASGLDHCDIDINECLEDRKSVCCORGDCI 851

```

```

QY 239 NQGTVCSCPGYLLDDNRSODINECEHNRHTCNLQGTCTYNLQGEKCI----- 290
DB 852 NTAGSTDCCTCPDGF-QLDNKTQDINECEHPG-LCGPQGECLNTEGSHFYCQGGFSIS 909
QY 291 -DPIRCEPEYLRLIS-----DN-----RCMC-----PAENPGCRDOPFTLLYRDM 328
DB 910 ADGRTCEIDECVNTWCDSHGFCDNTRAGSFRLCLYGGFQAQDQGGVD-----VNQC 963
QY 329 DVTSG 333
DB 964 ELLSG 968

```

Search completed: July 3, 2003, 18:23:14
 Job time : 10.6843 secs

Db 374 VNSPGSFRCCKGYGFDGJSRMCVDNRCORPGRLOGHCNCELTGLSCSVGFRLS 433
 Qy 175 EDRSCODVNECATENPCVOTCVNTYSGFICRDPGELEE-DGVHCSMDMCSF--SBE 231
 Db 434 VDRSCDINECSS-SPSCDCAVYGSYQCYCRGVLSDVDTCTEDIDECALPTGQH 492
 Qy 232 LCOHECVNPGYFVFCSP-PCYILLDDNRSCODINECHRNHTCMLOOTCVNLQSGFKCI 290
 Db 493 ICYRCININGSPSCSSGYSYRLAPNRCNDIDECVTGJHNCINTECFNIGCAFRL 552
 Qy 291 DPLRCEPFIARISDN-----RCMCPAENPCGROPFILYMDMVSGRSV----- 336
 Db 553 -AECPEENRRSATLQOERTDVRCI-----KSCRPDVYCVDPVHTISHVLSLPTF 606
 Qy 337 -----PADIFOMOA-TTRPG--AYYIFQIKSGNEGRE-----YMRQTPISATLVMTR 383
 Db 607 RETTREPFIETFLAITPPHPSAQNITFDITEGNIURDSFDIIRKYM--DGMTVGVVROVR 664
 Qy 384 PIKGPRIQDLEM-ITVNTVINFRGSSVIRLRIYVSQYF 423
 Db 665 PIVGPFAVAKLEKNYVGVGVSHR--NVNVRIFVSEYWF 703
 RESULT 14
 FBLL CAEEL STANDARD; PRT; 712 AA.
 ID FBLL CAEEL 077469; 077474; 095N23;
 AC 16-OCT-2001 (Rel. 40, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Fubulin-1 precursor.
 GN FBLLN OR F56H11.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodderinae; Caenorhabditis.
 NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99120531; PubMed=9923656;
 RA Barth J.L., Argaves K.M., Roark E.F., Little C.D., Argaves W.S.;
 RT Identification of chicken and C. elegans fibulin-1 homologs and
 RT characterization of the C. elegans fibulin-1 gene.";
 RL Matrix Biol. 17:635-646(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RA Lloyd C.R.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX (BY SIMILARITY).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; C (SHOWN HERE) AND D; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 9 EGF-LIKE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF051403; AAC28323.1; -
 DR EMBL; AF051403; AAC28324.1; -
 DR EMBL; AF051401; AAC28321.1; -
 DR EMBL; AF051402; AAC28322.1; -
 DR EMBL; Z68219; CAC35826.1; -
 DR EMBL; Z68749; CAC35826.1; JOINED.
 DR EMBL; Z68219; CAC35827.1; -
 DR EMBL; Z68749; CAC35827.1; JOINED.
 DR EMBL; Z68219; CAC35817.1; -
 DR EMBL; Z68749; CAC35817.1; JOINED.
 DR EMBL; Z68749; CAC35818.1; -

DR EMBL; Z68219; CAC35818.1; JOINED.
 DR WormRep; F56H11.1a; CE26701.
 DR HSSP; P16109; 1PSB.
 DR InterPro; IPR000020; Anaphylatoxin.
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF-Ca.
 DR Pfam; PF00008; EGF; 6.
 DR Pfam; PF01821; ANATO; 2.
 DR SMART; SM00104; ANATO; 2.
 DR SMART; SM00179; EGF-Ca; 6.
 DR SMART; SM00001; EGF-like; 3.
 DR PROSITE; PS00010; ASX_HYDROXYL; 4.
 DR PROSITE; PS01177; ANAPHYLATOXIN_L; 1.
 DR PROSITE; PS01166; EGF_2; 5.
 DR PROSITE; PS01187; EGF-Ca; 8.
 DR Signal; Glycoprotein; Extracellular matrix; Repeat; EGF-like domain;
 KW Calcium-binding; Alternative splicing.
 KW SIGNAL 1 17
 FT CHAIN 18 712
 FT DOMAIN 23 64
 FT DOMAIN 63 96
 FT DOMAIN 97 129
 FT DOMAIN 155 194
 FT DOMAIN 195 241
 FT DOMAIN 242 303
 FT DOMAIN 304 350
 FT DOMAIN 351 390
 FT DOMAIN 391 434
 FT DOMAIN 435 475
 FT DOMAIN 476 520
 FT DOMAIN 521 568
 FT DISULFID 23 49
 FT DISULFID 24 56
 FT DISULFID 37 94
 FT DISULFID 66 94
 FT DISULFID 79 95
 FT DISULFID 97 121
 FT DISULFID 98 128
 FT DISULFID 111 129
 FT DISULFID 159 168
 FT DISULFID 164 178
 FT DISULFID 180 193
 FT DISULFID 199 212
 FT DISULFID 206 221
 FT DISULFID 227 240
 FT DISULFID 246 259
 FT DISULFID 253 268
 FT DISULFID 291 302
 FT DISULFID 308 320
 FT DISULFID 314 329
 FT DISULFID 316 349
 FT DISULFID 335 365
 FT DISULFID 355 365
 FT DISULFID 360 374
 FT DISULFID 376 389
 FT DISULFID 395 409
 FT DISULFID 403 418
 FT DISULFID 420 433
 FT DISULFID 439 450
 FT DISULFID 446 459
 FT DISULFID 461 474
 FT DISULFID 480 495
 FT DISULFID 491 504
 FT DISULFID 506 519
 FT DISULFID 525 537
 FT DISULFID 530 546
 FT DISULFID 551 565
 FT VARSPLIC 564 712
 ANAPHYLATOXIN-LIKE 1.
 ANAPHYLATOXIN-LIKE 2.
 ANAPHYLATOXIN-LIKE 3.
 EGF-LIKE 1.
 EGF-LIKE 2.
 EGF-LIKE 3.
 EGF-LIKE 4.
 EGF-LIKE 5.
 EGF-LIKE 6.
 EGF-LIKE 7.
 EGF-LIKE 8.
 EGF-LIKE 9.
 EGF-LIKE 10.
 EGF-LIKE 11.
 EGF-LIKE 12.
 EGF-LIKE 13.
 EGF-LIKE 14.
 EGF-LIKE 15.
 EGF-LIKE 16.
 EGF-LIKE 17.
 EGF-LIKE 18.
 EGF-LIKE 19.
 EGF-LIKE 20.
 EGF-LIKE 21.
 EGF-LIKE 22.
 EGF-LIKE 23.
 EGF-LIKE 24.
 EGF-LIKE 25.
 EGF-LIKE 26.
 EGF-LIKE 27.
 EGF-LIKE 28.
 EGF-LIKE 29.
 EGF-LIKE 30.
 EGF-LIKE 31.
 EGF-LIKE 32.
 EGF-LIKE 33.
 EGF-LIKE 34.
 EGF-LIKE 35.
 EGF-LIKE 36.
 EGF-LIKE 37.
 EGF-LIKE 38.
 EGF-LIKE 39.
 EGF-LIKE 40.
 EGF-LIKE 41.
 EGF-LIKE 42.
 EGF-LIKE 43.
 EGF-LIKE 44.
 EGF-LIKE 45.
 EGF-LIKE 46.
 EGF-LIKE 47.
 EGF-LIKE 48.
 EGF-LIKE 49.
 EGF-LIKE 50.
 EGF-LIKE 51.
 EGF-LIKE 52.
 EGF-LIKE 53.
 EGF-LIKE 54.
 EGF-LIKE 55.
 EGF-LIKE 56.
 EGF-LIKE 57.
 EGF-LIKE 58.
 EGF-LIKE 59.
 EGF-LIKE 60.
 EGF-LIKE 61.
 EGF-LIKE 62.
 EGF-LIKE 63.
 EGF-LIKE 64.
 EGF-LIKE 65.
 EGF-LIKE 66.
 EGF-LIKE 67.
 EGF-LIKE 68.
 EGF-LIKE 69.
 EGF-LIKE 70.
 EGF-LIKE 71.
 EGF-LIKE 72.
 EGF-LIKE 73.
 EGF-LIKE 74.
 EGF-LIKE 75.
 EGF-LIKE 76.
 EGF-LIKE 77.
 EGF-LIKE 78.
 EGF-LIKE 79.
 EGF-LIKE 80.
 EGF-LIKE 81.
 EGF-LIKE 82.
 EGF-LIKE 83.
 EGF-LIKE 84.
 EGF-LIKE 85.
 EGF-LIKE 86.
 EGF-LIKE 87.
 EGF-LIKE 88.
 EGF-LIKE 89.
 EGF-LIKE 90.
 EGF-LIKE 91.
 EGF-LIKE 92.
 EGF-LIKE 93.
 EGF-LIKE 94.
 EGF-LIKE 95.
 EGF-LIKE 96.
 EGF-LIKE 97.
 EGF-LIKE 98.
 EGF-LIKE 99.
 EGF-LIKE 100.
 EGF-LIKE 101.
 EGF-LIKE 102.
 EGF-LIKE 103.
 EGF-LIKE 104.
 EGF-LIKE 105.
 EGF-LIKE 106.
 EGF-LIKE 107.
 EGF-LIKE 108.
 EGF-LIKE 109.
 EGF-LIKE 110.
 EGF-LIKE 111.
 EGF-LIKE 112.
 EGF-LIKE 113.
 EGF-LIKE 114.
 EGF-LIKE 115.
 EGF-LIKE 116.
 EGF-LIKE 117.
 EGF-LIKE 118.
 EGF-LIKE 119.
 EGF-LIKE 120.
 EGF-LIKE 121.
 EGF-LIKE 122.
 EGF-LIKE 123.
 EGF-LIKE 124.
 EGF-LIKE 125.
 EGF-LIKE 126.
 EGF-LIKE 127.
 EGF-LIKE 128.
 EGF-LIKE 129.
 EGF-LIKE 130.
 EGF-LIKE 131.
 EGF-LIKE 132.
 EGF-LIKE 133.
 EGF-LIKE 134.
 EGF-LIKE 135.
 EGF-LIKE 136.
 EGF-LIKE 137.
 EGF-LIKE 138.
 EGF-LIKE 139.
 EGF-LIKE 140.
 EGF-LIKE 141.
 EGF-LIKE 142.
 EGF-LIKE 143.
 EGF-LIKE 144.
 EGF-LIKE 145.
 EGF-LIKE 146.
 EGF-LIKE 147.
 EGF-LIKE 148.
 EGF-LIKE 149.
 EGF-LIKE 150.
 EGF-LIKE 151.
 EGF-LIKE 152.
 EGF-LIKE 153.
 EGF-LIKE 154.
 EGF-LIKE 155.
 EGF-LIKE 156.
 EGF-LIKE 157.
 EGF-LIKE 158.
 EGF-LIKE 159.
 EGF-LIKE 160.
 EGF-LIKE 161.
 EGF-LIKE 162.
 EGF-LIKE 163.
 EGF-LIKE 164.
 EGF-LIKE 165.
 EGF-LIKE 166.
 EGF-LIKE 167.
 EGF-LIKE 168.
 EGF-LIKE 169.
 EGF-LIKE 170.
 EGF-LIKE 171.
 EGF-LIKE 172.
 EGF-LIKE 173.
 EGF-LIKE 174.
 EGF-LIKE 175.
 EGF-LIKE 176.
 EGF-LIKE 177.
 EGF-LIKE 178.
 EGF-LIKE 179.
 EGF-LIKE 180.
 EGF-LIKE 181.
 EGF-LIKE 182.
 EGF-LIKE 183.
 EGF-LIKE 184.
 EGF-LIKE 185.
 EGF-LIKE 186.
 EGF-LIKE 187.
 EGF-LIKE 188.
 EGF-LIKE 189.
 EGF-LIKE 190.
 EGF-LIKE 191.
 EGF-LIKE 192.
 EGF-LIKE 193.
 EGF-LIKE 194.
 EGF-LIKE 195.
 EGF-LIKE 196.
 EGF-LIKE 197.
 EGF-LIKE 198.
 EGF-LIKE 199.
 EGF-LIKE 200.
 EGF-LIKE 201.
 EGF-LIKE 202.
 EGF-LIKE 203.
 EGF-LIKE 204.
 EGF-LIKE 205.
 EGF-LIKE 206.
 EGF-LIKE 207.
 EGF-LIKE 208.
 EGF-LIKE 209.
 EGF-LIKE 210.
 EGF-LIKE 211.
 EGF-LIKE 212.
 EGF-LIKE 213.
 EGF-LIKE 214.
 EGF-LIKE 215.
 EGF-LIKE 216.
 EGF-LIKE 217.
 EGF-LIKE 218.
 EGF-LIKE 219.
 EGF-LIKE 220.
 EGF-LIKE 221.
 EGF-LIKE 222.
 EGF-LIKE 223.
 EGF-LIKE 224.
 EGF-LIKE 225.
 EGF-LIKE 226.
 EGF-LIKE 227.
 EGF-LIKE 228.
 EGF-LIKE 229.
 EGF-LIKE 230.
 EGF-LIKE 231.
 EGF-LIKE 232.
 EGF-LIKE 233.
 EGF-LIKE 234.
 EGF-LIKE 235.
 EGF-LIKE 236.
 EGF-LIKE 237.
 EGF-LIKE 238.
 EGF-LIKE 239.
 EGF-LIKE 240.
 EGF-LIKE 241.
 EGF-LIKE 242.
 EGF-LIKE 243.
 EGF-LIKE 244.
 EGF-LIKE 245.
 EGF-LIKE 246.
 EGF-LIKE 247.
 EGF-LIKE 248.
 EGF-LIKE 249.
 EGF-LIKE 250.
 EGF-LIKE 251.
 EGF-LIKE 252.
 EGF-LIKE 253.
 EGF-LIKE 254.
 EGF-LIKE 255.
 EGF-LIKE 256.
 EGF-LIKE 257.
 EGF-LIKE 258.
 EGF-LIKE 259.
 EGF-LIKE 260.
 EGF-LIKE 261.
 EGF-LIKE 262.
 EGF-LIKE 263.
 EGF-LIKE 264.
 EGF-LIKE 265.
 EGF-LIKE 266.
 EGF-LIKE 267.
 EGF-LIKE 268.
 EGF-LIKE 269.
 EGF-LIKE 270.
 EGF-LIKE 271.
 EGF-LIKE 272.
 EGF-LIKE 273.
 EGF-LIKE 274.
 EGF-LIKE 275.
 EGF-LIKE 276.
 EGF-LIKE 277.
 EGF-LIKE 278.
 EGF-LIKE 279.
 EGF-LIKE 280.
 EGF-LIKE 281.
 EGF-LIKE 282.
 EGF-LIKE 283.
 EGF-LIKE 284.
 EGF-LIKE 285.
 EGF-LIKE 286.
 EGF-LIKE 287.
 EGF-LIKE 288.
 EGF-LIKE 289.
 EGF-LIKE 290.
 EGF-LIKE 291.
 EGF-LIKE 292.
 EGF-LIKE 293.
 EGF-LIKE 294.
 EGF-LIKE 295.
 EGF-LIKE 296.
 EGF-LIKE 297.
 EGF-LIKE 298.
 EGF-LIKE 299.
 EGF-LIKE 300.
 EGF-LIKE 301.
 EGF-LIKE 302.
 EGF-LIKE 303.
 EGF-LIKE 304.
 EGF-LIKE 305.
 EGF-LIKE 306.
 EGF-LIKE 307.
 EGF-LIKE 308.
 EGF-LIKE 309.
 EGF-LIKE 310.
 EGF-LIKE 311.
 EGF-LIKE 312.
 EGF-LIKE 313.
 EGF-LIKE 314.
 EGF-LIKE 315.
 EGF-LIKE 316.
 EGF-LIKE 317.
 EGF-LIKE 318.
 EGF-LIKE 319.
 EGF-LIKE 320.
 EGF-LIKE 321.
 EGF-LIKE 322.
 EGF-LIKE 323.
 EGF-LIKE 324.
 EGF-LIKE 325.
 EGF-LIKE 326.
 EGF-LIKE 327.
 EGF-LIKE 328.
 EGF-LIKE 329.
 EGF-LIKE 330.
 EGF-LIKE 331.
 EGF-LIKE 332.
 EGF-LIKE 333.
 EGF-LIKE 334.
 EGF-LIKE 335.
 EGF-LIKE 336.
 EGF-LIKE 337.
 EGF-LIKE 338.
 EGF-LIKE 339.
 EGF-LIKE 340.
 EGF-LIKE 341.
 EGF-LIKE 342.
 EGF-LIKE 343.
 EGF-LIKE 344.
 EGF-LIKE 345.
 EGF-LIKE 346.
 EGF-LIKE 347.
 EGF-LIKE 348.
 EGF-LIKE 349.
 EGF-LIKE 350.
 EGF-LIKE 351.
 EGF-LIKE 352.
 EGF-LIKE 353.
 EGF-LIKE 354.
 EGF-LIKE 355.
 EGF-LIKE 356.
 EGF-LIKE 357.
 EGF-LIKE 358.
 EGF-LIKE 359.
 EGF-LIKE 360.
 EGF-LIKE 361.
 EGF-LIKE 362.
 EGF-LIKE 363.
 EGF-LIKE 364.
 EGF-LIKE 365.
 EGF-LIKE 366.
 EGF-LIKE 367.
 EGF-LIKE 368.
 EGF-LIKE 369.
 EGF-LIKE 370.
 EGF-LIKE 371.
 EGF-LIKE 372.
 EGF-LIKE 373.
 EGF-LIKE 374.
 EGF-LIKE 375.
 EGF-LIKE 376.
 EGF-LIKE 377.
 EGF-LIKE 378.
 EGF-LIKE 379.
 EGF-LIKE 380.
 EGF-LIKE 381.
 EGF-LIKE 382.
 EGF-LIKE 383.
 EGF-LIKE 384.
 EGF-LIKE 385.
 EGF-LIKE 386.
 EGF-LIKE 387.
 EGF-LIKE 388.
 EGF-LIKE 389.
 EGF-LIKE 390.
 EGF-LIKE 391.
 EGF-LIKE 392.
 EGF-LIKE 393.
 EGF-LIKE 394.
 EGF-LIKE 395.
 EGF-LIKE 396.
 EGF-LIKE 397.
 EGF-LIKE 398.
 EGF-LIKE 399.
 EGF-LIKE 400.
 EGF-LIKE 401.
 EGF-LIKE 402.
 EGF-LIKE 403.
 EGF-LIKE 404.
 EGF-LIKE 405.
 EGF-LIKE 406.
 EGF-LIKE 407.
 EGF-LIKE 408.
 EGF-LIKE 409.
 EGF-LIKE 410.
 EGF-LIKE 411.
 EGF-LIKE 412.
 EGF-LIKE 413.
 EGF-LIKE 414.
 EGF-LIKE 415.
 EGF-LIKE 416.
 EGF-LIKE 417.
 EGF-LIKE 418.
 EGF-LIKE 419.
 EGF-LIKE 420.
 EGF-LIKE 421.
 EGF-LIKE 422.
 EGF-LIKE 423.
 EGF-LIKE 424.
 EGF-LIKE 425.
 EGF-LIKE 426.
 EGF-LIKE 427.
 EGF-LIKE 428.
 EGF-LIKE 429.
 EGF-LIKE 430.
 EGF-LIKE 431.
 EGF-LIKE 432.
 EGF-LIKE 433.
 EGF-LIKE 434.
 EGF-LIKE 435.
 EGF-LIKE 436.
 EGF-LIKE 437.
 EGF-LIKE 438.
 EGF-LIKE 439.
 EGF-LIKE 440.
 EGF-LIKE 441.
 EGF-LIKE 442.
 EGF-LIKE 443.
 EGF-LIKE 444.
 EGF-LIKE 445.
 EGF-LIKE 446.
 EGF-LIKE 447.
 EGF-LIKE 448.
 EGF-LIKE 449.
 EGF-LIKE 450.
 EGF-LIKE 451.
 EGF-LIKE 452.
 EGF-LIKE 453.
 EGF-LIKE 454.
 EGF-LIKE 455.
 EGF-LIKE 456.
 EGF-LIKE 457.
 EGF-LIKE 458.
 EGF-LIKE 459.
 EGF-LIKE 460.
 EGF-LIKE 461.
 EGF-LIKE 462.
 EGF-LIKE 463.
 EGF-LIKE 464.
 EGF-LIKE 465.
 EGF-LIKE 466.
 EGF-LIKE 467.
 EGF-LIKE 468.
 EGF-LIKE 469.
 EGF-LIKE 470.
 EGF-LIKE 471.
 EGF-LIKE 472.
 EGF-LIKE 473.
 EGF-LIKE 474.
 EGF-LIKE 475.
 EGF-LIKE 476.
 EGF-LIKE 477.
 EGF-LIKE 478.
 EGF-LIKE 479.
 EGF-LIKE 480.
 EGF-LIKE 481.
 EGF-LIKE 482.
 EGF-LIKE 483.
 EGF-LIKE 484.
 EGF-LIKE 485.
 EGF-LIKE 486.
 EGF-LIKE 487.
 EGF-LIKE 488.
 EGF-LIKE 489.
 EGF-LIKE 490.
 EGF-LIKE 491.
 EGF-LIKE 492.
 EGF-LIKE 493.
 EGF-LIKE 494.
 EGF-LIKE 495.
 EGF-LIKE 496.
 EGF-LIKE 497.
 EGF-LIKE 498.
 EGF-LIKE 499.
 EGF-LIKE 500.
 EGF-LIKE 501.
 EGF-LIKE 502.
 EGF-LIKE 503.
 EGF-LIKE 504.
 EGF-LIKE 505.
 EGF-LIKE 506.
 EGF-LIKE 507.
 EGF-LIKE 508.
 EGF-LIKE 509.
 EGF-LIKE 510.
 EGF-LIKE 511.
 EGF-LIKE 512.
 EGF-LIKE 513.
 EGF-LIKE 514.
 EGF-LIKE 515.
 EGF-LIKE 516.
 EGF-LIKE 517.
 EGF-LIKE 518.
 EGF-LIKE 519.
 EGF-LIKE 520.
 EGF-LIKE 521.
 EGF-LIKE 522.
 EGF-LIKE 523.
 EGF-LIKE 524.
 EGF-LIKE 525.
 EGF-LIKE 526.
 EGF-LIKE 527.
 EGF-LIKE 528.
 EGF-LIKE 529.
 EGF-LIKE 530.
 EGF-LIKE 531.
 EGF-LIKE 532.
 EGF-LIKE 533.
 EGF-LIKE 534.
 EGF-LIKE 535.
 EGF-LIKE 536.
 EGF-LIKE 537.
 EGF-LIKE 538.
 EGF-LIKE 539.
 EGF-LIKE 540.
 EGF-LIKE 541.
 EGF-LIKE 542.
 EGF-LIKE 543.
 EGF-LIKE 544.
 EGF-LIKE 545.
 EGF-LIKE 546.
 EGF-LIKE 547.
 EGF-LIKE 548.
 EGF-LIKE 549.
 EGF-LIKE 550.
 EGF-LIKE 551.
 EGF-LIKE 552.
 EGF-LIKE 553.
 EGF-LIKE 554.
 EGF-LIKE 555.
 EGF-LIKE 556.
 EGF-LIKE 557.
 EGF-LIKE 558.
 EGF-LIKE 559.
 EGF-LIKE 560.
 EGF-LIKE 561.
 EGF-LIKE 562.
 EGF-LIKE 563.
 EGF-LIKE 564.
 EGF-LIKE 565.
 EGF-LIKE 566.
 EGF-LIKE 567.
 EGF-LIKE 568.
 EGF-LIKE 569.
 EGF-LIKE 570.
 EGF-LIKE 571.
 EGF-LIKE 572.
 EGF-LIKE 573.
 EGF-LIKE 574.
 EGF-LIKE 575.
 EGF-LIKE 576.
 EGF-LIKE 577.
 EGF-LIKE 578.
 EGF-LIKE 579.
 EGF-LIKE 580.
 EGF-LIKE 581.
 EGF-LIKE 582.
 EGF-LIKE 583.
 EGF-LIKE 584.
 EGF-LIKE 585.
 EGF-LIKE 586.
 EGF-LIKE 587.
 EGF-LIKE 588.
 EGF-LIKE 589.
 EGF-LIKE 590.
 EGF-LIKE 591.
 EGF-LIKE 592.
 EGF-LIKE 593.
 EGF-LIKE 594.
 EGF-LIKE 595.
 EGF-LIKE 596.
 EGF-LIKE 597.
 EGF-LIKE 598.
 EGF-LIKE 599.
 EGF-LIKE 600.
 EGF-LIKE 601.
 EGF-LIKE 602.
 EGF-LIKE 603.
 EGF-LIKE 604.
 EGF-LIKE 605.
 EGF-LIKE 606.
 EGF-LIKE 607.
 EGF-LIKE 608.
 EGF-LIKE 609.
 EGF-LIKE 610.
 EGF-LIKE 611.
 EGF-LIKE 612.
 EGF-LIKE 613.
 EGF-LIKE 614.
 EGF-LIKE 615.
 EGF-LIKE 616.
 EGF-LIKE 617.
 EGF-LIKE 618.
 EGF-LIKE 619.
 EGF-LIKE 620.
 EGF-LIKE 621.
 EGF-LIKE 622.
 EGF-LIKE 623.
 EGF-LIKE 624.
 EGF-LIKE 625.
 EGF-LIKE 626.
 EGF-LIKE 627.
 EGF-LIKE 628.
 EGF-LIKE 629.
 EGF-LIKE 630.
 EGF-LIKE 631.
 EGF-LIKE 632.
 EGF-LIKE 633.
 EGF-LIKE 634.
 EGF-LIKE 635.
 EGF-LIKE 636.
 EGF-LIKE 637.
 EGF-LIKE 638.
 EGF-LIKE 639.
 EGF-LIKE 640.
 EGF-LIKE 641.
 EGF-LIKE 642.
 EGF-LIKE 643.
 EGF-LIKE 644.
 EGF-LIKE 645.
 EGF-LIKE 646.
 EGF-LIKE 647.
 EGF-LIKE 648.
 EGF-LIKE 649.
 EGF-LIKE 650.
 EGF-LIKE 651.
 EGF-LIKE 652.
 EGF-LIKE 653.
 EGF-LIKE 654.
 EGF-LIKE 655.
 EGF-LIKE 656.
 EGF-LIKE 657.
 EGF-LIKE 658.
 EGF-LIKE 659.
 EGF-LIKE 660.
 EGF-LIKE 661.
 EGF-LIKE 662.
 EGF-LIKE 663.
 EGF-LIKE 664.
 EGF-LIKE 665.
 EGF-LIKE 666.
 EGF-LIKE 667.
 EGF-LIKE 668.
 EGF-LIKE 669.
 EGF-LIKE 670.
 EGF-LIKE 671.
 EGF-LIKE 672.
 EGF-LIKE 673.
 EGF-LIKE 674.
 EGF-LIKE 675.
 EGF-LIKE 676.
 EGF-LIKE 677.
 EGF-LIKE 678.
 EGF-LIKE 679.
 EGF-LIKE 680.
 EGF-LIKE 681.
 EGF-LIKE 682.
 EGF-LIKE 683.
 EGF-LIKE 684.
 EGF-LIKE 685.
 EGF-LIKE 686.
 EGF-LIKE 687.
 EGF-LIKE 688.
 EGF-LIKE 689.
 EGF-LIKE 690.
 EGF-LIKE 691.
 EGF-LIKE 692.
 EGF-LIKE 693.
 EGF-LIKE 694.
 EGF-LIKE 695.
 EGF-LIKE 696.
 EGF-LIKE 697.
 EGF-LIKE 698.
 EGF-LIKE 699.
 EGF-LIKE 700.
 EGF-LIKE 701.
 EGF-LIKE 702.
 EGF-LIKE 703.
 EGF-LIKE 704.
 EGF-LIKE 705.
 EGF-LIKE 706.
 EGF-LIKE 707.
 EGF-LIKE 708.
 EGF-LIKE 709.
 EGF-LIKE 710.
 EGF-LIKE 711.
 EGF-LIKE 712.
 EGF-LIKE 713.
 EGF-LIKE 714.
 EGF-LIKE 715.
 EGF-LIKE 716.
 EGF-LIKE 717.
 EGF-LIKE 718.
 EGF-LIKE 719.
 EGF-LIKE 720.
 EGF-LIKE 721.
 EGF-LIKE 722.
 EGF-LIKE 723.
 EGF-LIKE 724.
 EGF-LIKE 725.
 EGF-LIKE 726.
 EGF-LIKE 727.
 EGF-LIKE 728.
 EGF-LIKE 729.
 EGF-LIKE 730.
 EGF-LIKE 731.
 EGF-LIKE 732.
 EGF-LIKE 733.
 EGF

[illegible]

FT	DISULFID	558	579		BY SIMILARITY.
FT	CARBOHYD	98	98		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	537	537		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	541	541		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARSPLIC	569	705		FROEKDTVRCIKSRPMDEACVRDPVHTVSHTVSLPTFR
FT					EETREBEIIFLAAPYPLYPANADIIFFITBSNLRSDIIT
FT					KRIEDGTAVGVROYRPIVGPYATVLAKEMNTVLLGGVSHR
FT					NNVNVAHIVSEIWF -> RCARPHEKNOECPRDLRTIYY
FT					HLSEPNIQVPAAVPRMGPSASVPDGSMLATAAGNEGGFF
FT					TTRKSHSGVVALKRPPEIPERDILLTVKMDLYRHGVTSSSF
FT					VAKLFIIVSAEL (IN ISOFORM D).
SQ	SEQUENCE	705 AA:	78056 MW;	FD3F06469A4BEE2B CRC64;	
	Query Match	29.2%;	Score 702;	DB 1; Length 705;	
	Best Local Similarity	35.2%;	Pred. No. 1.9e-46;		
	Matches 160;	Conservative 66;	Mismatches 154;	Indels 74;	Gaps 21.
QY	1 OCTNGFDLDROSGCCLDIDECRTPEACRGDMCYONGGYCICTPRTPVTKGPSINPYS	60			
Dd	295 QCKSKEFIDD-ALGNCIDINECISAPCPVGCGTINTGSIYVC-----	336			
QY	61 TPYSGPYAAPAPLSAPNPYPTISRRLICRFQGM-DENOCVDVDECATSHQCNPQTIC	119			
Dd	337 -----QKNVFN-----CGRGHLNEBSTRCVDVECAPPAEPCSKGHNC	375			
QY	120 INTEGGYTSCSTDGYWL--LEGOCHLIDEC-RY-GYCOOLCANPGSYSGTCNPGFTLN	174			
Dd	376 LNSPSPFCECKAGFYFDGISMTCDINMEGRPERLKGHCENPPGFHCSGSAFPILS	435			
QY	175 EDGNSCODVNECATLENPCVOICVNTYGSFICRDPGEYLEE-DGVHSGDMDECSF--SEF	231			
Dd	436 VDRGCEDEPNBEC-LNSPCSQBECANYGSGYCRCRGYLDSVDGTCEDIDECCALPTGCH	494			
QY	232 ICGHCVNQPGNYPCSCP-PGYILLDDNRSCODIRECHRNHTCMLOOTCVNLGGFCICI	290			
Dd	495 ICSTRCLINPSPFOCCSPSSGIRLPAPNRNOCODIDECTGTHNCSINTECFIOQSFKCL	554			
QY	291 DPICEEPYLRLISDN-----RCM--CPAENPGC-RDOPFTILYXMDMVVSGRSV--	336			
Dd	555 S-FECPENRYRSADTFREQKTIDYWCISCRRNDACVCRDPVHTVSHVISLPTREFTR	613			
QY	337 PADIFOMQATT-RIPG-AIYIFQIKSNEGHEPVM---ROIQPSATLVWTRPKGRE	390			
Dd	614 PEEIIFLAAYPLYPANADIIFFDTTECNTLDSFDIIKRYEDGMTVGVROYRPIVGPFY	673			
QY	391 IQOLDLEM-IIVNTVIINFEGSSVIRLRIVSOYPF	423			
Dd	674 AVLKLNNMYVLGGVYSHR--NVNVNHIVSEIWF	705			
RESULT 12					
ID	FBL2_HUMAN	STANDARD:	PRT:	1184 AA.	
AC	P98095:				
DT	01-NOV-1995 (Rel. 32, Created)				
DT	01-NOV-1995 (Rel. 32, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Fibulin-2 precursor.				
GN	FBLN2.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
OX	NCBI_Taxid=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Fibroblast;				
RX	MEDLINE=95104855; PubMed=7806230;				
RA	Zhang R.-Z., Pan T.-C., Zhang Z.-Y., Mattei M.-G., Timpl R.,				
RA	Chu M.-L.;				
RT	"Fibulin-2 (FBLN2): human cDNA sequence, mRNA expression, and mapping				
RT	of the gene on human and mouse chromosomes.";				
BC	Genomics 22:425-430(1994).				
LC	-1- FUNCTION=ITS BINDING TO FIBRONECTIN AND SOME OTHER LIGANDS IS				

[illegible]

FT	CONFLICT	140	159	HSGRKVAAGHVVHSSCAC -> TVANISIMPRRPILP
FT	CONFLICT	348	348	GF (IN REF. 2).
FT	CONFLICT	507	507	S -> L (IN REF. 2).
FT	CONFLICT	1102	1102	Q -> QO (IN REF. 2).
SO	SEQUENCE	1221 AA;	131818 MW;	87DB2A10A8FDC45F CRC64;
Query Match		30.3%;	Score 729.5;	DB 1;
Best Local Similarity		35.3%;	Pred. No.2.6e-48;	
Matches		145;	Conservative 59;	Mismatches 148;
			Indels 59;	Gaps 11;
QY	1	QCTNGFDLDRSGGCLDIDECRTIPEACRGDMCVNONGYLCLIPRTNPVYRGPSNPS	60	
Db	832	RCMGGF-LDPEGKNCVDINECTSLLEPCRSGFSCINTGSYTC-----	873	
QY	61	TPYSGPYPAAPPLSAPRYPIPSILICREBYQ-DESNOCVDYDECATSDHQCNPYIC	119	
Db	874	-----QRNPILVCGRGYHANESESECVNDECTGVHRCGEGLC	912	
QY	120	INTEGGYTCSTGDYV-LLEGQCLDIDECRYG---YCQCLANVPGSYSCCNPGFTLN	174	
Db	913	YNLPGSRCDCKPQFQADAPRGTIDVNECVASGRLCQHNCMPGPSRSCAAGFLA	972	
QY	175	EDGRSCDVNECATENPCVOTCVNTYGSFICRDPGLELEDGYHSCSDMDECS-FSEFLC	233	
Db	973	ADGHCEEDVNECEFTRR-CSQECANIYSGYCCYCRGOYLADHDGHTCTDIDECACGAGITC	10311	
QY	234	QHECVNPGPYFGSCP-PGYILLDNNSCODINCEHNHNCINLQPCYNLGGFKICDP	292	
Db	1032	TFRCVNVPGSIQCCAPEDGITMAANGSKDLDECALGTHNCSEAEYCHNIGSFRCL-R	1090	
QY	293	IRCEPYLRISDNRCMKPAENPGCD-----QFTLLYRMDVYSGRSPADIFOMAT	346	
Db	1091	FDCPPNYVRSQTKC---ERTQODLIECQTSFARLHYOLNQTGLLVPAHIFRLGPA	1146	
QY	347	TRYGAYVYIPIKSGNGERFEMOTGPISATLWTRPKRPREIOLDLEM	397	
Db	1147	PAPGDTIISLTITKNGNEGYVTRLNAYTVSLSQNSVLEPRPADDVM	1197	
RESULT 11				
FBLL_MOUSE				
ID	FBLL_MOUSE	STANDARD;	PRT:	705 AA.
AC	Q08879; Q08878;			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	01-OCT-1994 (Rel. 30, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Fibulin-1 precursor (Basement-membrane protein 90) (BM-90).			
GN	FBLN1.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID:10090;			
RN	[1]			
CC	SEQUENCE FROM N.A.			
RC	STRAIN-MK31.			
RX	MEDLINE-9335887; PubMed-8354280.			
RA	Pan T.-C., Kluge M., Zhang R.Z., Mayer U., Timpi R., Chu M.-L.,			
RT	"Sequence of extracellular mouse protein BM-90/fibulin and its			
RT	calcium-dependent binding to other basement-membrane ligands."			
RL	Eur. J. Biochem. 215:733-740(1993).			
CC	-1 SUBCELLULAR LOCATION: Secreted; extracellular matrix.			
CC	-1 ALTERNATIVE PRODUCTS: 4 ISOFORMS: A, B, C AND D (SHOWN HERE); ARE			
CC	PRODUCED BY ALTERNATIVE SPLICING. THEY DIFFER ONLY IN THEIR C-			
CC	TERMINAL REGIONS.			
CC	-1 SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.			
CC	-1 SIMILARITY: CONTAINS 9 EGF-LIKE DOMAINS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			

```

FT DISULFID 89 108 BY SIMILARITY.
FT DISULFID 110 134 BY SIMILARITY.
FT DISULFID 111 141 BY SIMILARITY.
FT DISULFID 124 142 BY SIMILARITY.
FT DISULFID 181 191 BY SIMILARITY.
FT DISULFID 187 200 BY SIMILARITY.
FT DISULFID 202 215 BY SIMILARITY.
FT DISULFID 221 234 BY SIMILARITY.
FT DISULFID 228 243 BY SIMILARITY.
FT DISULFID 249 261 BY SIMILARITY.
FT DISULFID 267 280 BY SIMILARITY.
FT DISULFID 274 289 BY SIMILARITY.
FT DISULFID 295 307 BY SIMILARITY.
FT DISULFID 313 326 BY SIMILARITY.
FT DISULFID 320 335 BY SIMILARITY.
FT DISULFID 342 355 BY SIMILARITY.
FT DISULFID 361 374 BY SIMILARITY.
FT DISULFID 368 383 BY SIMILARITY.
FT DISULFID 385 398 BY SIMILARITY.
FT DISULFID 404 416 BY SIMILARITY.
FT DISULFID 412 425 BY SIMILARITY.
FT DISULFID 427 440 BY SIMILARITY.
FT DISULFID 446 455 BY SIMILARITY.
FT DISULFID 451 464 BY SIMILARITY.
FT DISULFID 466 480 BY SIMILARITY.
FT DISULFID 486 499 BY SIMILARITY.
FT DISULFID 495 508 BY SIMILARITY.
FT DISULFID 510 524 BY SIMILARITY.
FT DISULFID 530 543 BY SIMILARITY.
FT DISULFID 537 552 BY SIMILARITY.
FT DISULFID 557 569 BY SIMILARITY.
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 536 536 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 540 540 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 638 638 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 684 AA: 75623 MW: 163807A094739199 CMC64;

```

```

Query Match 32.0%; Score 769.5; DB 1; Length 684;
Best Local Similarity 36.5%; Pred. No. 1.3e-51;
Matches 159; Conservative 62; Mismatches 150; Indels 65; Gaps 15;

```

```

QY 1 QCNNGDLDNROSCCLDIDCRITPEACRQDMCMVONGYLCTIPRTNPNVRYGPNYS 60
DB 294 QCNNGDLDNROSCCLDIDCRITPEACRQDMCMVONGYLCTIPRTNPNVRYGPNYS 60
QY 61 TPYSGPYPAAPPLAPNPVTISRLPLCRFGYQWDE-SNOCVYDDECATDSHCNPFOIC 119
DB 341 -----SCGRGYHLNEDGTRCVADVDECSSDQPCGEGHVC 374
QY 120 INTEGYTSCDIDGW--LLEGQCLDIDECRY--GYCOQLCANVPSSCTCNPGETLN 174
DB 375 INGPNGYRCCKSGSFDVISRFCIDINECRARRGRCAHKCENTPSSYCTCMGFSLTS 434
QY 175 EGRSGODVNECACTENPCVQVNTGSEFCRCRDPGELEE-DGVHCSMDCEGSEF--SEF 231
DB 435 SGRSGEDNECES-SPCSOECANVGSYCCYCRGQQLDIDGISEDEDECALPTGSH 493
QY 232 LCOHECVNPGTFCSCP-PGYTILDNFRSCODINECHNRHTNLOTCVNLGGFRCI 290
DB 494 ICSFRICINPGSQCTCPSTGYRLAPNARNCODIDEVAETHNCSFNETCNIIDGGRCL 553
QY 291 DTRCEEPRLRISDNRC-MCPA-ENPGCRDQPTIILYRMDVYSGRSVPADIFQMAQYTR 348
DB 554 S-LCEPENRKSDTRCERLPCNENKCOQLPLRTFYHLSFPNINQVPTDIFPMGNSNA 612
QY 349 YPGAYVYIPIKSGNENREFYMRGTGPIISALVMTPTPKGREGLODLEMT-----YNTV 403
DB 613 VPGDXILLISIGNOGCFPTTKKVNHNHSGIVVMOROTTEPRDLITOMOLTRHGTVNTE 672
QY 404 INFRGSSVYLRIRIYVS 419
DB 673 I-----AKLFVFS 681

```

```

RESULT 10
FBI2_MOUSE STANDARD; PRT; 1221 AA.
AC P37889; O9WU12;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibulin-2 precursor.
GN FBI2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-35.
RC TISSUE=Fibroblast;
RX MEDLINE=94064787; PubMed=8245130;
RA Pan T.-C., Sasaki T., Zhang R.-Z., Faessler R., Timpi R., Chu M.-L.;
RT "Structure and expression of fibulin-2, a novel extracellular matrix
RT protein with multiple EGF-like repeats and consensus motifs for
RT calcium binding.";
RL J. Cell Biol. 123:1269-1277(1993).
RN [2]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=99337686; PubMed=10406956;
RA Graessl S., Sticot F.-X., Gotta S., Chu M.-L.;
RT "Mouse fibulin-2 gene. Complete exon-intron organization and promoter
RT characterization.";
RL Eur. J. Biochem. 263:471-477(1999).
CC -1- FUNCTION: ITS BINDING TO FIBRONECTIN AND SOME OTHER LIGANDS IS
CC -1- CALCIUM DEPENDENT.
CC -1- SUBUNIT: HOMOTRIMER; DISULFIDE-LINKED.
CC -1- SUBCELLULAR LOCATION: Extracellular matrix.
CC -1- ALTERNATIVE PRODUCTS: At least 2 isoforms; 1 (shown here) and
CC 2/EGF3-less; are produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: COMPONENT OF BOTH BASEMENT MEMBRANES AND OTHER
CC CONNECTIVE TISSUES.
CC -1- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 11 EGF-LIKE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DB EMBL; X75285; CAA53040.1; -
DR EMBL; AF135253; AAD34456.1; -
DR EMBL; AF135239; AAD34456.1; JOINED.
DR EMBL; AF135240; AAD34456.1; JOINED.
DR EMBL; AF135241; AAD34456.1; JOINED.
DR EMBL; AF135242; AAD34456.1; JOINED.
DR EMBL; AF135243; AAD34456.1; JOINED.
DR EMBL; AF135244; AAD34456.1; JOINED.
DR EMBL; AF135245; AAD34456.1; JOINED.
DR EMBL; AF135246; AAD34456.1; JOINED.
DR EMBL; AF135247; AAD34456.1; JOINED.
DR EMBL; AF135248; AAD34456.1; JOINED.
DR EMBL; AF135249; AAD34456.1; JOINED.
DR EMBL; AF135250; AAD34456.1; JOINED.
DR EMBL; AF135251; AAD34456.1; JOINED.
DR EMBL; AF135252; AAD34456.1; JOINED.
DR PIR; A49457; A49457.
DR HSSP; P00736; IAPQ.
DR MGD; MGI:95488; FBI2.
DR InterPro; IPR000020; Anaphylatoxin.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR Pfam; PF00008; EGF; 6.

```

[illegible]

DB		444	PVSANMLVKSLSGPREHIVDLEMLTVSSIGTFRRSVLRLLIIVGPSEF	493
RESULT_9				
ID	FBL1_CHICK	STANDARD:	PRT:	684 AA.
AC	073775;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Fibulin-1 precursor.			
GN	FBLN1.			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauiria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
CC	Gallus.			
CX	NCBI_TaxID=9031;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	Bioinformatics-99120531; PubMed=9923656;			
RA	Martin J.L., Argreaves K.M., Roark E.F., Little C.D., Argreaves W.S.;			
RT	"Identification of chicken and C. elegans fibulin-1 homologs and			
RT	characterization of the C. elegans fibulin-1 gene.";			
RL	Matrix Biol. 17:635-646(1998).			
CC	-1- SUBCELLULAR LOCATION: Secreted; extracellular matrix.			
CC	-1- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.			
CC	-1- SIMILARITY: CONTAINS 9 EGF-LIKE DOMAINS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AF051400; AAC05388.1; --			
DR	HSSP; P00742; IHCG.			
DR	InterPro; IPR0000020; Anaphylatoxin.			
DR	InterPro; IPR000152; Asx_hydroxyl.			
DR	InterPro; IPR000561; EGF-like.			
DR	InterPro; IPR001881; EGF_Ca.			
DR	Pfam; PF00008; EGF_6			
DR	Pfam; PF01821; ANATO_2.			
DR	SMART; SMART0104; ANATO; 3.			
DR	SMART; SMART0179; EGF_CA; 8.			
DR	SMART; SMO0001; EGF_like; 1.			
DR	PROSITE; PS00010; ASX_HYDROXYL; 5.			
DR	PROSITE; PS01177; ANAPHYLATOXIN_1; 1.			
DR	PROSITE; PS01178; ANAPHYLATOXIN_2; 2.			
DR	PROSITE; PS00022; EGF_1; FALSE_NEG.			
DR	PROSITE; PS01186; EGF_2; 3.			
DR	PROSITE; PS01187; EGF_CA; 8.			
KW	Signal; Glycoprotein; Extracellular matrix; Repeat; EGF-like domain;			
KW	Calcium-binding.			
FT	SIGNAL	1..25	POTENTIAL.	
FT	CHAIN	26..684	FIBULIN-1.	
FT	DOMAIN	33..74	ANAPHYLATOXIN-LIKE 1.	
FT	DOMAIN	75..109	ANAPHYLATOXIN-LIKE 2.	
FT	DOMAIN	110..142	ANAPHYLATOXIN-LIKE 3.	
FT	DOMAIN	177..216	EGF-LIKE 1.	
FT	DOMAIN	217..262	EGF-LIKE 2.	
FT	DOMAIN	263..308	EGF-LIKE 3.	
FT	DOMAIN	309..356	EGF-LIKE 4.	
FT	DOMAIN	357..399	EGF-LIKE 5.	
FT	DOMAIN	400..441	EGF-LIKE 6.	
FT	DOMAIN	442..481	EGF-LIKE 7.	
FT	DOMAIN	482..525	EGF-LIKE 8.	
FT	DOMAIN	526..570	EGF-LIKE 9.	
FT	DISULFID	33..59	BY SIMILARITY.	
FT	DISULFID	34..66	BY SIMILARITY.	
FT	DISULFID	47..67	BY SIMILARITY.	
FT	DISULFID	76..107	BY SIMILARITY.	

```

DR PROSITE: PS0010; ASX_HYDROXYL; 4.
DR PROSITE: PS0022; EGF_1; FALSE_NEG.
DR PROSITE: PS0186; EGF_2; 4.
DR PROSITE: PS0187; EGF_CA; 6.
KW Repeat; EGF-like domain; Calcium-binding; Glycoprotein; Signal.
FT SIGNAL 1 17
FT CHAIN 18 493
FT DOMAIN 26 71
FT DOMAIN 173 213
FT DOMAIN 214 253
FT DOMAIN 254 293
FT DOMAIN 294 333
FT DOMAIN 334 378
FT DISULFID 177 190
FT DISULFID 184 199
FT DISULFID 201 212
FT DISULFID 218 228
FT DISULFID 224 237
FT DISULFID 239 252
FT DISULFID 258 268
FT DISULFID 264 277
FT DISULFID 279 292
FT DISULFID 298 309
FT DISULFID 305 318
FT DISULFID 320 332
FT DISULFID 338 350
FT DISULFID 344 359
FT DISULFID 365 377
SQ SEQUENCE 493 AA; 54596 MW; 22DAFD70BACFICA5 CRC64;

Query Match 46.1%; Score 1108.5; DB 1; Length 493;
Best Local Similarity 43.6%; Pred. No. 1.2e-77;
Matches 204; Conservative 68; Mismatches 149; Indels 47; Gaps 5;

QY 1 OCTGFDIDROSGGCLDIDECRTIPACRGDMCMVONGSYLCIPRTNPPYRGPSNPYS 60
DB 28 OCTGGEYMDPYRQOCDEIDEDVIDPACGKGMKVHNGYGLCLPRTAQIIVNNEQPOE 87
QY 61 TPYS-----GPPAARPLSLAPNPT----- 81
DB 88 TPAAEASGAATGTIAARSMATSGVTPGGGFTASATAVAGPEVOTGRNNEFVIRNPAPQ 147
QY 82 -----ISRLPFCRGYQMDENOCVDVDECATDSHOCNPTQICINTEGYTSCSTGQYWL 136
DB 148 RPSNPSHRIRGICAGYEGSEHNVCODIDECTSGTHNCRIDQVCINLRGSGTCLRGYOK 207
QY 137 LEGGCLDIDECRY-GYCOQLCANVPGSYSCGTPGFTLNEDGRSCDVNECATENPCVQT 195
DB 208 RGEQCVDIDECSPVPCYHQGCVNTPGSPFCQCPGQFLAANNVTCVDINECDASNCAQ 267
QY 196 CVNTGYSFICRCGDELEEDGCVHSCDMECSFSEFLCOHECVNOCGYFECSPRGYIL 255
DB 268 CYNILGSTCCNOCQYELSDRLNCEIDECRTSYLCQYQCVNEGKFSQCPQGYOV 327
QY 256 DDNKSCDINECEHNTCNLOQTCYNLGGFKCIDPICEBEYRLISDNRCMCPAENPG 315
DB 328 -RSTGCDINCEFTNE-CREDECMNHYHGRFCYRPNQCDPYPVLTSENRCYCPSNM 365
QY 316 CRDOPFTILYRDMVDVSGVPADIFQOMATRYRGAYYIFQIKSGNEGREFYRQGTPT 375
DB 386 CRDVPQSVIVKYNMIRSDRSVPDIFQIQAATTIVANTINTPRIKSGNENGEFYLRQTSV 445
QY 376 SATLVMTRPIKRGPREIQLDLEMTIVNTYINFRGSSVIRLRIYSQPF 423
DB 446 SAMLVKSLGTGPREHIVGLEMLVSSIGTFRTSVLRLLIIVGPSF 493

RESULT 8
FBL3_HUMAN STANDARD; PRT; 493 AA.
AC Q12805;
DT 16-OCT-2001 (Rel. 40, Created)

```

```

DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE EGF-containing fibulin-like extracellular matrix protein 1 precursor
DE (Fibulin-3) (FBL-3) (Fibrillin-like protein) (Extracellular protein
DE S1-5).
GN EFEMP1 OR FBLN3 OR FBNL
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A., AND POSSIBLE ALTERNATIVE SPLICING.
RC MEDLINE=95097983; PubMed=7799918;
RA Lecka-Czenik B., Lumpkin C.K. Jr., Goldstein S.;
RT "An overexpressed gene transcript in senescent and quiescent human
RT fibroblasts encoding a novel protein in the epidermal growth factor-
RT like repeat family stimulates DNA synthesis.";
RL Mol. Cell. Biol. 15:120-128(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97001163; PubMed=8812496;
RA Ikegawa S., Toda T., Okui K., Nakamura Y.;
RT "Structure and chromosomal assignment of the human S1-5 gene (FBLN)
RT that is highly homologous to fibrillin.";
RL Genomics 35:590-592(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20068041; PubMed=10601734;
RA Giltay R., Timpl R., Kostka G.;
RT "Sequence, recombinant expression and tissue localization of two novel
RT extracellular matrix proteins, fibulin-3 and fibulin-4.";
RL Matrix Biol. 18:469-480(1999).
RN [4]
RP VARIANT DHRD/MLVT TRP-345, AND VARIANT PHE-220.
RX MEDLINE=99295941; PubMed=10369267;
RA Stone E.M., Lotery A.J., Munier F.L., Heon E., Pignet B., Guymer R.H.,
RA Vandenburgh K., Cousin P., Nishimura D., Swiderski R.E., Silvestri G.,
RA Mackey D.A., Hagerman G.S., Bird A.C., Sheffield V.C.,
RA Schorderet D.F.;
RT "A single EFEMP1 mutation associated with both malattia Leventinese
RT and Doyme honeycomb retinal dystrophy.";
RL Nat. Genet. 22:199-202(1999).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- ALTERNATIVE PRODUCTS: 4 isoforms; 1 (shown here), 2, 3 and 4; may
CC be produced by alternative splicing.
CC -1- DISEASE: DEFECTS IN EFEMP1 ARE A CAUSE OF DOYNE HONEYCOMB RETINAL
CC DYSTROPHY (DHRD) ALSO KNOWN AS MALATTIA LEVENTINESE (MLVT OR ML),
CC AN AUTOSOMAL DOMINANT DISEASE CHARACTERIZED BY YELLOW-WHITE
CC DEPOSITS KNOWN AS DRUSEN THAT ACCUMULATE BENEATH THE RETINAL
CC PIGMENT EPITHELIUM.
CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U03877; AAA65590.1; -.
DR HSSP: P35555; 1EMN.
DR Genew: HGNC:3218; EFEMP1.
DR MIM: 601548; -.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_CA.
DR Pfam: PF00008; EGF; 3.
DR SMART: SM00179; EGF_CA; 4.
DR SMART: SM00001; EGF-like; 2.
DR PROSITE: PS0010; ASX_HYDROXYL; 4.

```

Ox	Mammalia,Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Rn	[1]
Rp	SEQUENCE FROM N.A.
Rc	STRAIN=C57BL/6J;
Rx	MEDLINE=99308589; PubMed=10380882;
Ra	Gallagher W.M., Argentini M., Sierra V., Bracco L., Debussche L.,
Rt	Conseller E.;
Rt	"MMP1: a novel mutant p53-specific protein partner with oncogenic
Rt	properties.";
Rl	Oncogene 18:3608-3616(1999).
Cc	-1- SUBUNIT: BINDS PREFERENTIALLY TO P53 MUTANTS.
Cc	-1- SUBCELLULAR LOCATION: Secreted.
Cc	-1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
Cc	-----
Cc	This SWISS-PROT entry is copyright. It is produced through a collaboration
Cc	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
Cc	the European Bioinformatics Institute. There are no restrictions on its
Cc	use by non-profit institutions as long as its content is in no way
Cc	modified and this statement is not removed. Usage by and for commercial
Cc	entities requires a license agreement (See http://www.isb-sib.ch/announce/
Cc	or send an email to licenselsb-sib.ch).
Cc	-----
Dc	EMBL; AF104223; AAd45219.1; -
Dc	HSSP; P00736; IAPQ.
Dc	MGI; MG1:1891209; Efemp2.
Dc	InterPro: IPRO00152; Asx_hydroxy1.
Dc	InterPro: IPRO00561; EGF-like.
Dc	InterPro: IPRO01881; EGF-Ca.
Dc	InterPro: IPRO01491; Thrmomodulin.
Dc	Pfam; Pf00008; EGF; 4.
Dc	PRINTS; PRO0907; THRMOMODULN.
Dc	SMART; SMO0179; EGF_CA; 4.
Dc	SMART; SMO0001; EGF_Like; 2.
Dc	PROSITE; PS00022; ASX_HYDROXYL; 4.
Dc	PROSITE; PS01186; EGF_L; FALSE_NEG.
Dc	PROSITE; PS01187; EGF_CA; 6.
Kw	Repeat; EGF-like domain; Calcium-binding; Glycoprotein; Signal.
Ft	SIGNAL 1 25 POTENTIAL.
Ft	CHAIN 26 443 EGF-CONTAINING FIBULIN-LIKE EXTRACELLULAR
Ft	MATRIX PROTEIN 2.
Ft	EGF-Like 1, DIVERGENT.
Ft	EGF-Like 2, CALCIUM-BINDING (POTENTIAL).
Ft	EGF-Like 3, CALCIUM-BINDING (POTENTIAL).
Ft	EGF-Like 4, CALCIUM-BINDING (POTENTIAL).
Ft	EGF-Like 5, CALCIUM-BINDING (POTENTIAL).
Ft	EGF-Like 6, CALCIUM-BINDING (POTENTIAL).
Ft	BY SIMILARITY.
Ft	BY SIMILARITY.
Ft	BY SIMILARITY.
Ft	BY SIMILARITY.
Ft	BY SIMILARITY.
Ft	BY SIMILARITY.
Ft	BY SIMILARITY.
Ft	BY SIMILARITY.
Ft	BY SIMILARITY.
Ft	BY SIMILARITY.
Ft	BY SIMILARITY.
Ft	BY SIMILARITY.
Ft	BY SIMILARITY.
Ft	BY SIMILARITY.
Ft	BY SIMILARITY.
Ft	N-LINKED (GLCNAC. .) (POTENTIAL).
Ft	N-LINKED (GLCNAC. .) (POTENTIAL).
Ft	CARBONYD 198 198
Ft	CARBONYD 394 394
Ft	SEQUENCE 443 AA; 49425 MW; 4969C0328A23DD8 CRC64;
Query Match	52.98; Score 1272.5; DB 1; Length 443;
Best Local Similarity	52.08; Pred. NO. 3.2e-90;
Matches 220; Conservative	65; Mismatches 121; Indels 17; Gaps 3;

[illegible]

ID	FBL3_RAT	STANDARD;	PRT;	493 AA.
AC	035568;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	EGF-containing fibulin-like extracellular matrix protein 1 precursor (Fibulin-3) (FBLN-3) (716 protein).			
GN	EFEMP1 OR FBLN3.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID:10116;			
RA	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lung;			
RX	MEDLINE=97415782; PubMed=9268694;			
RA	Ozeki T., Kondo K., Nakamura Y., Ichimiya S., Nakagawara A.,			
RT	Sakiyama S.;			
RL	"Interaction of D441, a DAN-binding protein, with the epidermal growth factor-like protein, S(1-5)."			
RJ	Biochem. Biophys. Res. Commun. 237:245-250(1997).			
CC	-1- SUBCELLULAR LOCATION: Secreted.			
CC	-1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on the use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; D89730; BAA22265.1; .			
DR	HSSP; P35555; IEMN.			
DR	InterPro; IPR000152; Asx_hydroxyl.			
DR	InterPro; IPR000561; EGF-like.			
DR	InterPro; IPR001881; EGF_Ca.			
DR	Pfam; PF00008; EGF_3.			
DR	SMART; SM00179; EGF_CA_4.			
DR	SMART; SM00001; EGF_like_2.			

DT 16-OCT-2001 (Rel. 40, last annotation update)
 DE EGF-containing fibulin-like extracellular matrix protein 2 precursor
 DE (Fibulin-4) (FibL-4) (OPH1 protein).
 GN EFEMP2 OR FBLL4.
 OC Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Melanoma;
 RX MEDLINE=20068041; PubMed=10601734;
 RX Giltay R., Timpl R., Kostka G.;
 RT "Sequence, recombinant expression and tissue localization of two novel
 RT extracellular matrix proteins, fibulin-3 and fibulin-4.";
 RL Matrix Biol. 18:469-480(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC Zemel R., Shaul Y.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20435063; PubMed=10992184;
 RX Katsanis N., Venable S., Smith J.R., Lupski J.R.;
 RT "Isolation of a paralog of the Doyme honeycomb retinal dystrophy gene
 RT from the multiple retinopathy critical region on 11q13.";
 RL Hum. Genet. 106:66-72(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RL Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).
 CC -----
 CC EMBL: AJ132819; CA10791.2; -
 DR EMBL: AF093119; AAC62108.1; -
 DR EMBL: AF109121; AAF65188.1; -
 DR EMBL: BC010456; AAH10456.1; -
 DR HSSP: P35555; 1EMN.
 DR Genew: HGNC:3219; EFEMP2.
 DR MIM: 604633; -
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR001491; Thrombomodulin.
 DR Pfam: PF00008; EGF_4.
 DR PRINTS: PRO0907; THROMBOMODULN.
 DR SMART: SM00179; EGF_CA; 4.
 DR SMART: SM00001; EGF_Like; 2.
 DR PROSITE: PS00010; ASX_HYDROXYL; 4.
 DR PROSITE: PS00022; EGF_1; FALSE_NEG.
 DR PROSITE: PS01186; EGF_2; 4.
 DR PROSITE: PS01187; EGF_CA; 6.
 KW Repeat; EGF-like domain; Calcium-binding; Glycoprotein; Signal.
 FT SIGNAL 1 25
 FT CHAIN 26 443
 FT DOMAIN 36 81
 FT DOMAIN 123 163
 FT DOMAIN 164 202
 FT DOMAIN 203 242
 FT DOMAIN 243 282
 FT DOMAIN 283 328
 FT DISULFID 127 140
 EGF-CONTAINING FIBULIN-LIKE EXTRACELLULAR
 MATRIX PROTEIN 2.
 EGF-LIKE 1, DIVERGENT.
 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
 BY SIMILARITY.

FT DISULFID 134 149
 FT DISULFID 151 162
 FT DISULFID 168 177
 FT DISULFID 173 186
 FT DISULFID 188 201
 FT DISULFID 207 217
 FT DISULFID 213 226
 FT DISULFID 228 241
 FT DISULFID 247 258
 FT DISULFID 254 267
 FT DISULFID 269 281
 FT DISULFID 287 300
 FT DISULFID 294 309
 FT DISULFID 315 327
 FT CARBOHYD 198 198
 FT CARBOHYD 394 394
 FT CONFLICT 5 5
 FT CONFLICT 44 51
 FT CONFLICT 103 111
 FT CONFLICT 294 294
 FT CONFLICT 354 356
 FT CONFLICT 355 355
 SO SEQUENCE 443 AA; 49391 MW; 9E9AC2393780D3B8 CRC64;
 Query Match 53.0%; Score 1276.5; DB 1; Length 443;
 Best Local Similarity 52.0%; Pred. No. 1,6e-90;
 Matches 220; Conservative 67; Mismatches 119; Indels 17; Gaps 4;
 QY 1 OCTNGEDLDROSGGOCIDDECRTPACGDDMMCVNONGYLCIDPTNPYRGPSPYIS 60
 DB ECTDTEVEMPDSDOCHCDVNECLTIPACGEMKCIHHYGGYLCIPSAVAINDLG--- 93
 QY 61 TPYSGPYPAADPLSAPNYPITISRLPICRFGYQMDSESNOCVDYDECATDSHOCNPTQICI 120
 DB 94 ---EGP-PEPVPAQHFN-----PCPPGVPEPPDQDSCVDYDECAQALHDCRPQDCH 141
 QY 121 NTEGGYTOSCTGYMLLEBQCLDIDECRIGYQOQLCANPAGYSTCNGFTLNEDGRSC 180
 DB 142 NLPGSYOCTCPGYKRIKGEVVIDECRRYVNOHRVNLPGSFRCCEGFGOLGPNRSC 201
 QY 181 ODVNECATENPCVQTCVMTVYSGFICRCDPGELEEDGVHCSMDDECSFSEFLCOHCVNO 240
 DB 202 VDVNECDMGAPCEGECFNSYGFELCRCHGEGELHRGFSQSDIDESYSYLQYCVNE 261
 QY 241 PGTYRSCPPGYTILLDDNRSQDINECEHRNHTCNLOQTCVNLGGFKCIDPICEPYL 300
 DB 262 PGRFSCHCPQGYQLL-ATRLCODIDECESGAHQCSAOTCVNFHGGYRCVDFNRCEPYI 320
 QY 301 RISDRKCPAENPGCRDOPFTILYRDMVYVGRSVAPDIPQMAOTRRPGAYITQIIS 360
 DB 321 QVSENRCCLPASNPLCRDPPSSIVHRMTITERSVPADVFOIQATSVPGAVNAFOIRA 380
 QY 361 GNEGRFYMRTGPISATLVMTPTIKGPEIOLDLEMTIVNVINRGSSVILRTIYVSG 420
 DB 381 GNSQGDFTYRQINNVASMLYLARPYTGPREVYLDLEMTVMNLSMYSRASVLRITVFGA 440
 QY 421 YPF 423
 DB 441 YTF 443
 RESULT 6
 FBLL_MOUSE
 ID FBLL_MOUSE STANDARD: PRT; 443 AA.
 AC Q9WV03;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, last sequence update)
 DT 16-OCT-2001 (Rel. 40, last annotation update)
 DE EGF-containing fibulin-like extracellular matrix protein 2 precursor
 DE (Fibulin-4) (FibL-4) (mutant p53 binding protein 1).
 GN EFEMP2 OR FBLL4 OR MBPL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Db 26 QCTNGFDLDROSGCCLDIDECRTIPEACRGDMCMVONGYLICTPRTNPNRYGPNPYS 85
QY 61 TPYSGPAPAPPLASAPNYPTISRLICRFGYOMDESQCVYDDECATDSHOCNPTQICI 120
Db 86 TTSISGPAPAPPLASAPNYPTISRLICRFGYOMDESQCVYDDECATDSHOCNPTQICI 145
QY 121 MTEGGYTCTSDYGLWLEGGCLDIDECRYGYCOQLCANVPSSYSTCNPNGFTLNEDGRSC 180
Db 146 MTEGGYTCTSDYGLWLEGGCLDIDECRYGYCOQLCANVPSSYSTCNPNGFTLNEDGRSC 205
QY 181 QDVNECATENPCVQTCVNTYGSFICRCDPGYLEEDGVHCDMDCSFSEFLCQHECVNQ 240
Db 206 QDVNECATENPCVQTCVNTYGSFICRCDPGYLEEDGVHCDMDCSFSEFLCQHECVNQ 265
QY 241 PGTYFSCPPGYILLDNRSQCQDINECEHRNHTCNLAQTCYNLAGFRCIDPICEEPTL 300
Db 266 PGTYFSCPPGYILLDNRSQCQDINECEHRNHTCTSLQTCYNLAGFRCIDPICEEPTL 325
QY 301 RISDNRCMCPAENPCRDQPTILYRDMVYSGNSVPADIFOMATTRYPGAYYIFQIKS 360
Db 326 LIGENRCMCPAENPCRDQPTILYRDMVYSGNSVPADIFOMATTRYPGAYYIFQIKS 385
QY 361 GNGREPYMROTPISATLVWTRPIKGRPREIQLDLEMTVTVMVNFRRSSVRLRLIYVSQ 420
Db 386 GNGREPYMROTPISATLVWTRPIKGRPREIQLDLEMTVTVMVNFRRSSVRLRLIYVSQ 445
QY 421 YPF 423
Db 446 YPF 448

```

RESULT 4
FBL4_CRIGR STANDARD; PRT; 443 AA.

```

AC 055058;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE EGF-containing fibulin-like extracellular matrix protein 2 precursor
GN (Fibulin-4) (FBLN-4) (H411 protein).
OS EGFMP2 OR FBLN4.
OS Crictetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Crictetinae;
OC Crictetulus.
OX NCBI_Taxid=10029;
RN [1]
RP SEQUENCE FROM N.A.
RT TISSUE=Ovary;
RL Heine H., Delude R.L., Monks B., Golenbock D.T.;
   Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF046870; AAC03101.1; -
DR HSSP: P00736; IAPQ.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR Pfam: PF00008; EGF_4.
DR SMART: SM00179; EGF_CA; 4.
DR SMART: SM00001; EGF_Like; 2.
DR PROSITE: PS00010; ASX_HYDROXYL; 4.
DR PROSITE: PS00022; EGF_1; FALSE_NMG.
DR PROSITE: PS01186; EGF_2; 4.

```

```

DR PROSITE: PS01187; EGF_CA; 6.
KW Repeat; EGF-like domain; Calcium-binding; Glycoprotein; signal.
FT SIGNL 1 25
FT CHAIN 26 443
FT DOMAIN 36 81
FT DOMAIN 123 163
FT DOMAIN 164 202
FT DOMAIN 203 242
FT DOMAIN 243 282
FT DOMAIN 283 328
FT DISULFID 127 140
FT DISULFID 134 149
FT DISULFID 151 162
FT DISULFID 168 177
FT DISULFID 173 186
FT DISULFID 188 201
FT DISULFID 207 217
FT DISULFID 213 226
FT DISULFID 228 241
FT DISULFID 247 258
FT DISULFID 254 267
FT DISULFID 269 281
FT DISULFID 287 300
FT DISULFID 294 309
FT DISULFID 315 327
FT CARBOHYD 198 198
FT CARBOHYD 394 394
SQ SEQUENCE 443 AA; 49432 MW; 0BCFE5D73239E5F CRC64;

```

Query Match 53.08; Score 1276.5; DB 1; Length 443;
Best Local Similarity 52.28; Pred. No. 1.6e-90;
Matches 221; Conservative 64; Mismatches 121; Indels 17; Gaps 4;

```

QY 1 QCTNGFDLDROSGCCLDIDECRTIPEACRGDMCMVONGYLICTPRTNPNRYGPNPYS 60
Db 38 ECTDGYEMDADSOHCRCRVNDECLTPEACRGDMKCNHGYGLCLPRAVAINDLHG---- 93
QY 61 TPYSGPAPAPPLASAPNYPTISRLICRFGYOMDESQCVYDDECATDSHOCNPTQICI 120
Db 94 ---EGP-PPVYPPAQRHN-----PCPPYEDEDSQCVYDDECATDSHOCNPTQICI 141
QY 121 MTEGGYTCTSDYGLWLEGGCLDIDECRYGYCOQLCANVPSSYSTCNPNGFTLNEDGRSC 180
Db 142 NLPGSYQCTCPDGYRKYRGPCEVDIDECRYRQCQRHCVMLPSFRCQCEPGFQIGPNNSC 201
QY 181 QDVNECATENPCVQTCVNTYGSFICRCDPGYLEEDGVHCDMDCSFSEFLCQHECVNQ 240
Db 202 VDVNECDMGAPCEQRCFNSYGTFLCRQNGYELHHDGFCSDIDECSSYSLQYRCVNE 261
QY 241 PGTYFSCPPGYILLDNRSQCQDINECEHRNHTCNLAQTCYNLAGFRCIDPICEEPTL 300
Db 262 PGTYFSCPPGYILLDNRSQCQDINECEHRNHTCNLAQTCYNLAGFRCIDPICEEPTL 320
QY 301 RISDNRCMCPAENPCRDQPTILYRDMVYSGNSVPADIFOMATTRYPGAYYIFQIKS 360
Db 321 QVSDNRCFCVSNPLCRQPSIYHRYMSITSESVADVFQIQATSYVPGAYNAFQIARA 380
QY 361 GNGREPYMROTPISATLVWTRPIKGRPREIQLDLEMTVTVMVNFRRSSVRLRLIYVSQ 420
Db 381 GNGREPYMROTPISATLVWTRPIKGRPREIQLDLEMTVTVMVNFRRSSVRLRLIYVSQ 440
QY 421 YPF 423
Db 441 YTF 443

```

RESULT 5
FBL4_HUMAN STANDARD; PRT; 443 AA.
ID FBL4_HUMAN
AC 095967; O75967;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)


```

FT DOMAIN 247 287 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 288 333 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT SITE 54 56 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 131 144 BY SIMILARITY.
FT DISULFID 136 153 BY SIMILARITY.
FT DISULFID 153 166 BY SIMILARITY.
FT DISULFID 172 181 BY SIMILARITY.
FT DISULFID 177 190 BY SIMILARITY.
FT DISULFID 192 205 BY SIMILARITY.
FT DISULFID 211 221 BY SIMILARITY.
FT DISULFID 217 230 BY SIMILARITY.
FT DISULFID 232 245 BY SIMILARITY.
FT DISULFID 251 262 BY SIMILARITY.
FT DISULFID 273 286 BY SIMILARITY.
FT DISULFID 292 305 BY SIMILARITY.
FT DISULFID 299 314 BY SIMILARITY.
FT DISULFID 320 332 BY SIMILARITY.
FT CARBOHYD 283 283 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 238 238 L -> P (IN REF. 2).
SO SEQUENCE 448 AA; 50160 MW; E6BC68F7B14B714 CRC64;

Query Match 95.7%; Score 2303; DB 1; Length 448;
Best Local Similarity 94.8%; Pred. No. 6,6e-169;
Matches 401; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

OY 1 OCTNGFDLDRSGCCDIDECRTIPKACRGDMKCVNNGGYLCIPRTNPYRGSPYSPYS 60
DB 26 OCTNGFDLDRSGCCDIDECRTIPKACRGDMKCVNNGGYLCIPRTNPYRGSPYSPYS 85
OY 61 TPYSGPYPAAPPLSAPNYPTTISRPLICRGYOMDSNOCVDDECATDSHCNPTQICI 120
DB 86 TSYSGPYPAAPPLSAPNYPTTISRPLICRGYOMDSNOCVDDECATDSHCNPTQICI 145
OY 121 NTEGGYTCCTGTYWLLBEGCCDIDECRTYGCQOLCANVPGSSYCTCNPGFTLNEDGRSC 180
DB 146 NTEGGYTCCTGTYWLLBEGCCDIDECRTYGCQOLCANVPGSSYCTCNPGFTLNEDGRSC 205
OY 181 ODVNECATNPGCVOTCVNTYGSFICRDPGELEBDSVHCSMDNDECSFSEFLCOHECVNO 240
DB 206 ODVNECATNPGCVOTCVNTYGSFICRDPGELEBDSVHCSMDNDECSFSEFLCOHECVNO 265
OY 241 PGTYGSCSPGYLLDNDNRSCODINECEHNRHTCNLOQTCTYNLQGGFKCIDPICEPEPYL 300
DB 266 PGTYGSCSPGYLLDNDNRSCODINECEHNRHTCTPLQCTYNLQGGFKCIDPICEPEPYL 325
OY 301 RLSNRCMCPAENPCGRDOPFTLLYRDMDVYSGRSVPADIFOMQATTRYPGAYYIFQIKS 360
DB 326 LIGDNRCMCPAENPCGRDOPFTLLYRDMDVYSGRSVPADIFOMQATTRYPGAYYIFQIKS 385
OY 361 GNEGREFYRORTGPISATLVMPRTIKGPREIOLDELTNTYINENGSSVIRLRIYVSQ 420
DB 386 GNEGREFYRORTGPISATLVMPRTIKGPREIOLDELTNTYINENGSSVIRLRIYVSQ 445
OY 421 YPF 423
DB 446 YPF 448

RESULT 3
FBL5_MOUSE STANDARD; PRT; 448 AA.
AC 09MVH9:
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DE F1bulin-5 precursor (F1bul-5) (Developmental arteries and neural crest
DE EGF-like protein) (Dance).
GN FBLN5 OR DANCE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

```

OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99357779; PubMed=10428923;
RA Nakamura T., Ruiz-Lozano P., Lindner V., Yabe D., Taniwaki M.,
RA Furukawa Y., Kobuke K., Tashiro K., Lu Z., Andon N.L., Schaub R.,
RA Matsumori A., Sasayama S., Chien K.R., Honjo T.;
RA "DANCE, a novel secreted RGD protein expressed in developing,
RA atherosclerotic, and balloon-injured arteries.";
RL J. Biol. Chem. 274:22476-22483(1999).
CC -1- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH
CC INTERACTION OF INTEGRINS AND THE RGD MOTIF. COULD BE A VASCULAR
CC LIGAND FOR INTEGRIN RECEPTORS AND MAY PLAY A ROLE IN VASCULAR
CC DEVELOPMENT AND REMODELING.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF112151; AM041767.1; -.
DR HSP: P00736; IAP0.
DR MGD: MGI:1346091; Fbln5.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR Pfam: PF00008; EGF: 4.
DR SMART: SM00179; EGF_CA: 4.
DR SMART: SM00001; EGF_Like: 2.
DR PROSITE: PS00010; ASX_HYDROXYL: 4.
DR PROSITE: PS00022; EGF_1; FALSE_NEG.
DR PROSITE: PS01186; EGF_2; 4.
DR PROSITE: PS01187; EGF_CA: 6.
KM Repeat: EGF-like domain; Calcium-binding; Glycoprotein; Signal.
FT SIGNAL 1 23
FT CHAIN 24 448
FT DOMAIN 24 69 EGF-LIKE 1, DIVERGENT.
FT DOMAIN 127 167 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 168 206 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 207 246 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 247 287 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 288 333 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT SITE 333 356 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 131 144 BY SIMILARITY.
FT DISULFID 138 153 BY SIMILARITY.
FT DISULFID 155 166 BY SIMILARITY.
FT DISULFID 172 181 BY SIMILARITY.
FT DISULFID 177 190 BY SIMILARITY.
FT DISULFID 192 205 BY SIMILARITY.
FT DISULFID 211 221 BY SIMILARITY.
FT DISULFID 217 230 BY SIMILARITY.
FT DISULFID 232 245 BY SIMILARITY.
FT DISULFID 251 262 BY SIMILARITY.
FT DISULFID 258 271 BY SIMILARITY.
FT DISULFID 273 286 BY SIMILARITY.
FT DISULFID 292 305 BY SIMILARITY.
FT DISULFID 299 314 BY SIMILARITY.
FT DISULFID 320 332 BY SIMILARITY.
FT CARBOHYD 283 283 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 448 AA; 50193 MW; F15CC70CCBFDF97 CRC64;

Query Match 95.6%; Score 2302; DB 1; Length 448;
Best Local Similarity 94.8%; Pred. No. 7,9e-169;
Matches 401; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

OY 1 OCTNGFDLDRSGCCDIDECRTIPKACRGDMKCVNNGGYLCIPRTNPYRGSPYSPYS 60
|||||
```

DR Genew; HGNC:3602; FBLS.
 DR MIM; 604580; -
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF_Ca.
 DR Pfam: PF00008; EGF; 4.
 DR SMART; SM00179; EGF_CA; 4.
 DR SMART; SM00001; EGF_Like; 2.
 DR PROSITE; PS00010; ASX_HYDROXYL; 4.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 4.
 DR PROSITE; PS01187; EGF_CA; 6.
 DR Repeat; EGF-like domain; Calcium-binding; glycoprotein; signal.
 KW SIGNAL.
 FT CHAIN 1 23
 FT SIGNAL 1 23
 FT DOMAIN 24 448
 FT DOMAIN 127 69
 FT DOMAIN 168 206
 FT DOMAIN 207 246
 FT DOMAIN 247 287
 FT DOMAIN 288 333
 FT SITE 54 56
 FT DISULFID 131 144
 FT DISULFID 138 153
 FT DISULFID 155 166
 FT DISULFID 172 181
 FT DISULFID 177 190
 FT DISULFID 192 205
 FT DISULFID 211 221
 FT DISULFID 217 230
 FT DISULFID 232 245
 FT DISULFID 251 262
 FT DISULFID 273 286
 FT DISULFID 278 292
 FT DISULFID 299 314
 FT DISULFID 320 332
 FT CARBOHYD 283 283
 FT CARBOHYD 296 296
 FT CONFLICT 69 70
 FT CONFLICT 147 148
 FT CONFLICT 147 148
 SQ SEQUENCE 448 AA; 50180 MW; 19FCAS1FDA328003 CRC64;

Query Match 100.0%; Score 2407; DB 1; Length 448;
 Best Local Similarity 100.0%; Pred. No. 7.5e-177;
 Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 QCTNGFPLDRSGGCLDIDECRTIEACRGDMCVNONGYLCIPRTNVPYRGPSNPYS 60
 26 QCTNGFPLDRSGGCLDIDECRTIEACRGDMCVNONGYLCIPRTNVPYRGPSNPYS 85
 61 TPYSGPAPAAPPTISAPNYPTISRPILICFRGYOMDESNOVCVDECATDSHOCNPOICI 120
 86 TPYSGPAPAAPPTISAPNYPTISRPILICFRGYOMDESNOVCVDECATDSHOCNPOICI 145
 121 NTEGGYTCSTGYWLLGGCLDIDECRYGYCOQCANVGSYSCYCNPGFTLNEDGRSC 180
 146 NTEGGYTCSTGYWLLGGCLDIDECRYGYCOQCANVGSYSCYCNPGFTLNEDGRSC 205
 181 QDVNECATENPCVOTCVNTYSFICRCPCGYLEEDGVCHSCSMDECSFSEFLQHCVCNO 240
 206 QDVNECATENPCVOTCVNTYSFICRCPCGYLEEDGVCHSCSMDECSFSEFLQHCVCNO 265
 241 PGTYFCSCPPIYLLDNRSCODINECHRNHTCNLOQTCYNLGGFKCIDPRCEPPT 300
 266 PGTYFCSCPPIYLLDNRSCODINECHRNHTCNLOQTCYNLGGFKCIDPRCEPPT 325
 301 RISDNRCMCPAENGCRCQDPTIILYRDMDVYSGRSVPADIFOMQATRTYRGAYYIFQIKS 360
 326 RISDNRCMCPAENGCRCQDPTIILYRDMDVYSGRSVPADIFOMQATRTYRGAYYIFQIKS 385
 361 GNEGREYMRQTGTSATLVWTRPIKGPKEIOLDLEMTVNTVINFRGSSVIRLRIYVSQ 420

DB 386 GNEGREYMRQTGTSATLVWTRPIKGPKEIOLDLEMTVNTVINFRGSSVIRLRIYVSQ 445
 OY 421 YPF 423
 DB 446 YPF 448

RESULT 2
 FBLS_RAT
 ID FBLS_RAT STANDARD: PRT: 448 AA.
 AC Q9WVH8; Q9R284;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Fibulin-5 precursor (FBLS-5) (Developmental arteries and neural crest
 DE EGF-like protein) (Dance) (Embryonic vascular EGF repeat-containing
 DE protein) (EVEC).
 GN FBLS OR DANCE.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99357779; PubMed=10428823;
 RA Nakamura T., Ruiz-Lozano P., Lindner V., Yabe D., Tanikawa M.,
 RA Furukawa Y., Kobike K., Tashiro K., Lu Z., Andon N.L., Schaub R.,
 RA Matsunori A., Sasayama S., Chien K.R., Honjo T.;
 RT "DANCE, a novel secreted RGD protein expressed in developing,
 RT atherosclerotic, and balloon-injured arteries";
 RL J. Biol. Chem. 274:22476-22483(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99278197; PubMed=10347091;
 RA Kowal R.C., Richardson J.A., Miano J.M., Olson E.N.;
 RT "EVEC, a novel epidermal growth factor-like repeat-containing protein
 RT upregulated in embryonic and diseased adult vasculature";
 RL Circ. Res. 84:1166-1176(1999).
 CC -1- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH
 CC INTERACTION OF INTEGRINS AND THE RGD MOTIF. COULD BE A VASCULAR
 CC LIGAND FOR INTEGRIN RECEPTORS AND MAY PLAY A ROLE IN VASCULAR
 CC DEVELOPMENT AND REMODELING.
 CC -1- SUBCELLULAR LOCATION: secreted.
 CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF112153; AAD41769.1; -;
 DR EMBL; AF137350; AAD25101.1; -;
 DR HSSP; P00736; IAPQ.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF_Ca.
 DR Pfam; PF00008; EGF; 4.
 DR SMART; SM00179; EGF_CA; 4.
 DR SMART; SM00001; EGF_Like; 2.
 DR PROSITE; PS00010; ASX_HYDROXYL; 4.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 4.
 DR PROSITE; PS01187; EGF_CA; 6.
 DR Repeat; EGF-like domain; Calcium-binding; glycoprotein; signal.
 KW SIGNAL.
 FT CHAIN 1 23
 FT SIGNAL 1 23
 FT DOMAIN 24 448
 FT DOMAIN 127 69
 FT DOMAIN 168 206
 FT DOMAIN 207 246

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 17:40:07 ; Search time 10.6843 Seconds
(without alignments)
1642.086 Million cell updates/sec

Title: US-09-674-379A-14

Perfect score: 2407
Sequence: 1 OCTNGFDLDROSGOCLDIDE.....INFGSSVRLRLVYSQPF 423

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2407	100.0	448	1	FBL5_HUMAN
2	2303	95.7	448	1	FBL5_RAT
3	2302	95.6	448	1	FBL5_MOUSE
4	1276.5	53.0	443	1	FBL4_CRIGR
5	1276.5	53.0	443	1	FBL4_HUMAN
6	1272.5	52.9	443	1	FBL4_MOUSE
7	1108.5	46.1	493	1	FBL3_RAT
8	1093.5	45.4	493	1	FBL3_HUMAN
9	769.5	32.0	684	1	FBL1_CHICK
10	729.5	30.3	1221	1	FBL1_MOUSE
11	702	29.2	705	1	FBL2_HUMAN
12	701.5	29.1	1184	1	FBL2_MOUSE
13	694	28.8	703	1	FBL1_CAEEL
14	596	24.8	712	1	FBL1_MOUSE
15	539	22.4	1394	1	FBL1_HUMAN
16	539	22.4	1595	1	FBL1_MOUSE
17	523.5	21.7	2871	1	FBN1_HUMAN
18	522.5	21.7	2871	1	FBN1_BOVIN
19	518	21.5	1712	1	FBN1_RAT
20	516.5	21.5	2871	1	FBN1_PIG
21	509.5	21.2	2871	1	FBN1_MOUSE
22	509.5	21.2	2907	1	FBN2_MOUSE
23	508.5	21.1	2911	1	FBN2_HUMAN
24	455	18.9	956	1	MTN2_HUMAN
25	455	18.9	956	1	MTN2_MOUSE
26	431	17.9	956	1	MTN2_MOUSE
27	392	16.3	886	1	EMR1_MOUSE
28	375	15.6	2470	1	EMR1_HUMAN
29	372	15.5	810	1	NEIL2_MOUSE
30	372	15.5	810	1	NEIL2_HUMAN
31	369	15.3	2471	1	NEIL2_MOUSE
32	367	15.2	2471	1	NEIL2_HUMAN
33	364	15.1	810	1	NEIL1_RAT

34	361.5	15.0	816	1	NEIL2_HUMAN	Q99435 homo sapien
35	357.5	14.9	816	1	NEIL2_RAT	O62918 rattus norv
36	355.5	14.8	816	1	NEIL_CHICK	O90827 gallus gall
37	353.5	14.7	1964	1	NTC4_MOUSE	P31695 mus musculu
38	351.5	14.6	2437	1	NTC1_BRARE	P46530 diacydantio
39	350.5	14.6	652	1	CD93_HUMAN	Q9npv3 homo sapien
40	348.5	14.5	2703	1	NOTC_DROME	P07207 drosophila
41	347.5	14.4	644	1	CD93_MOUSE	O89103 mus musculu
42	345.5	14.4	2556	1	NTC1_HUMAN	P46531 homo sapien
43	340	14.1	2524	1	NOTC_XENTLA	P21783 xenopus lae
44	337	14.0	835	1	CD97_HUMAN	P48960 homo sapien
45	335.5	13.9	2531	1	NTC1_RAT	Q07008 rattus norv

ALIGNMENTS

RESULT 1
FBL5_HUMAN STANDARD; PRT; 448 AA.

AC Q99435: 075966;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fblulin-5 precursor (FBL-5) (Developmental arteries and neural crest
DE EGF-like protein) (Dance) (Urine p50 protein) (UP50).
GN FBLN5 OR DANCE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Carnivora; Homiidae; Homo.
OX NCBI_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RA Kostka G., (MAR-1999) to the EMBL/Genbank/DBJ databases.
RL Submitted
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE:99357779; PubMed:10428823;
RA Nakamura T., Ruiz-Lozano P., Lindner V., Yabe D., Tanikawa M.,
RA Furukawa Y., Kobuke K., Tashiro K., Lu Z., Andon N.L., Schaub R.,
RA Matsumori A., Sasayama S., Chien K.R., Honjo T.;
RT "DANCE, a novel secreted RGD protein expressed in developing,
RT atherosclerotic, and balloon-injured arteries".
RL J. Biol. Chem. 274:22476-22483(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=urine;
RA Zemel R., Sholto O., Shaul Y.;
RL Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH
CC INTERACTION OF INTEGRINS AND THE RGD MOTIF. COULD BE A VASCULAR
CC LIGAND FOR INTEGRIN RECEPTORS AND MAY PLAY A ROLE IN VASCULAR
CC DEVELOPMENT AND REMODELING.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN HEART, OVARY, AND
CC COLON BUT ALSO IN KIDNEY, PANCREAS, TESTIS, LUNG, AND PLACENTA.
CC NOT DETECTABLE IN BRAIN, LIVER, THYMUS, PROSTATE, OR PERIPHERAL
CC BLOOD LEUCOCYTES.
CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AJ133490; CAB38568.1; -
CC EMBL: AF112152; AAD41768.1; -
CC EMBL: AF093118; AAC62107.1; -
CC HSSP: P00736; IAP0.

RESULT 15

A55567

fibrillin 1 - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 20-Feb-1995 #sequence, revision 20-Feb-1995 #text_change 02-Aug-2002

C:Accession: A55567

R:Titstra, D.J.; Li, L.; Potter, K.A.; Womack, J.; Byers, P.H.

Genomics 23, 480-485, 1994

A:Title: Sequence of the coding region of the bovine fibrillin cDNA and localization to

A:Reference number: A55567; MUID:95137597; PMID:7835900

A:Accession: A55567

A:Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-2871 <TIL>

C:Cross-references: GB:I28748; NID:9508427; PIDN:AAA74122.1; PID:9508428

C:Superfamily: fibrillin 1; EGF homology

F:I201-1236/Domain: EGF homology <EGF>

Query Match

21.7%; Score 522.5; DB 2; Length 2871;

Best Local Similarity 29.9%; Pred. No. 1,1e-26;

Matches 120; Conservative 45; Mismatches 124; Indels 113; Gaps 16;

```
QY      1 OCTNGEDLDROSGQCLIDECRTIPEACRGDMCVNONGYLCLPTNPVYRGPSNPYS 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      1054 RCDSEGFALDSEERCTDIDECRISPDLC-GRGQCVNTPGDFEC--KDEGYESGF----- 1105
QY      61 TPYSGPYPAAPPLSAFNPYISR-----PLICRFGYQMDSE-----NQ 99
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      1106 -----MMKNCMDIDECQDPDLICRGVCLNTEGSRCECPGHQAPNISA 1152
QY      100 CVDVDECATDSHOCNPTQICINTEGTYCSTDGWLEGO--CLDIDECRY--GYCOOL 155
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      1153 CIDINECELTAHLIC-PHGRVCNLIQKQACACNPGYHSTPDLFCVDIDECSTIMNGCETP 1211
QY      156 CANVPGSYSTCNCNPFNLNEDGRSCQDVNECATENP----- 191
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      1212 CTNSEGSYECSCQPGFALMPDRSCTDIDEC-EDNPNICDGGQCTNIPGEYRCLCYDGP 1270
QY      192 -----CVO-----TCVNTYGSFTCRCDPGYLEEDGVNCSMDDECSFSE 230
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      1271 ASEDMKTCTVYNNECDLNPNICLSTGTCENTKGSFICHDMGYSGKKGTGCTDINECEIGA 1330
QY      231 FLCQHE--CVNPGTYFCSPPGYILLDNRSQDINECEHRNHTCNLQOTCYNLGGPK 288
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      1331 HNCDRHAYCTNTAGSFRCSCSPGWI--GDGIKCTDLDSCSNGTNHCQHADCKNTMGSYR 1388
QY      289 CIDPIRCEEPYL-----RISDNRCMCPAENPGCRDOP 320
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      1389 CL-----CKEGYTGDFGTCTDLDCESENLILC--GNGQCINAP 1424
```

Search completed: July 3, 2003, 18:22:42
Job time : 17.9977 secs

A:Cross-references: GB:M34057; NID:9339547; PIDN:AAA61160.1; PID:9339548
 C:Superfamily: unassigned EGF-related proteins; EGF homology
 C:Keywords: alternative splicing
 F:750-791/Domain: EGF homology <EGF>

Query Match 22.4%; Score 539; DB 2; Length 1394;
 Best Local Similarity 29.6%; Pred. No. 4,66-28;

Matches 126; Conservative 51; Mismatches 126; Indels 122; Gaps 17;

QY 2 CTGPFDRSGGCLDDECRITPACRGDMCVNONGYLCIPRTNPNVNGPSNPYST 61
 DB 573 CTEGYRFEQQRCCVDDICTQVHLC-SQGRCENTESFLCI----- 614
 QY 62 PYSGPYAAPPLSAPNYPTISR-----CRFGYOMESNOCVVD 104
 DB 615 -CPAGFMASBEENICIDVDDECLRPDVGEGHCVTVGAFRCYECDSGRMRGRCEID 673
 QY 105 ECATDSHQCNPTQICINTEGGYTC-SCTDGYMLEGQCLDDEC-RVGYCOO-ICANYPG 161
 DB 674 ECLNPS-TC-PDEQCVNPSGSCVPCTEGFRGNNGCLDDECLPNCANNGCSNLEG 731
 QY 162 SYSCGNPFTLNEDEGSCDVNVECATENECVQ----- 194
 DB 732 STWCSCHKGTRTPDRHNCRDIDECQCGNLCVNGCKNTGESFRCTCGGYQLSNAKDC 791
 QY 195 -----TCVNTYGSFICRDPGELEEDGHCSDMDECSFSELCOH-ECV 238
 DB 792 EDIDECQHRLCAHGCRNTEGSEFCQCDGYSAGSLDCHDEDINECLERKSVQCRGDCI 851
 QY 239 NQPGYFCSCPPGYILLDNRSQCDINECEHRNHTCNLOQTCYNLQGGFKCI----- 290
 DB 852 NTAQSYDCCTCPDPE-GLDMKTCODINECEHPG-LCSPQEGCLNTEGSEFHCQCGFSIS 909
 QY 291 -DPIRCEEPYLRIS-----DN-----RCMC-----PANPCGRQDPFTILYDM 328
 DB 910 AGRCTEDIDECVNNVNCDSHGRCNDNTAGSFRCCLCYGQAPQDQGCVD-----VNEC 963
 QY 329 DVSNG 333
 DB 964 ELLSG 968

RESULT 14

A47221
 fibrillin 1 precursor - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 02-Jun-1995 #sequence, revision 25-Apr-1997 #text, change 02-Aug-2002
 C:Accession: A47221; S15355; S17064; I59574; S17062; S62111; A34198
 R:Corson, G.M.; Chabberg, S.C.; Dietz, H.C.; Charbonneau, N.L.; Sakai, L.Y.
 Genomics 17, 476-484, 1993
 A:Title: Fibrillin binds calcium and is coded by cDNAs that reveal a multidomain structure
 A:Reference number: A47221; MUID:94010947; PMID:7691719
 A:Accession: A47221
 A:Molecule type: mRNA
 A:Residues: 1-337, 'T', 339-1029 <COR>
 A:Cross-references: GB:X63556
 R:Perle, L.V.; D'Alessio, M.; Ramirez, F.; Lynch, J.; Sykes, B.; Panglman, T.; Bonad
 Hum. Mol. Genet. 2, 961-968, 1993
 A:Title: Genomic organization of the sequence coding for fibrillin, the defective gene F
 A:Reference number: I54355; MUID:93372860; PMID:8364578
 A:Accession: I54355
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 132-3002 <PER>
 A:Cross-references: GB:11923; NID:9306745; PIDN:AA02036.1; PID:9306746
 R:Maalen, C.L.; Corson, G.M.; Maddox, B.K.; Glanville, R.W.; Sakai, L.Y.
 Nature 352, 334-337, 1991
 A:Title: Partial sequence of a candidate gene for the Marfan syndrome.
 A:Reference number: S17064; MUID:91304568; PMID:1852207
 A:Accession: S17064
 A:Molecule type: mRNA
 A:Residues: 1030-3002 <MAS>
 A:Cross-references: EMBL:X63556

R,Dietz, H.C.; Valle, D.; Franccomano, C.A.; Kendzior, R.J.
 Science 259, 680-683, 1993
 A:Title: The skipping of constitutive exons in vivo induced by nonsense mutations.
 A:Reference number: I59574; MUID:93157831; PMID:8430317
 A:Accession: I59574

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 2217-2288, 'T', 2290-2325 <RES>

A:Cross-references: GB:S54426; NID:9264860; PIDN:AA025244.1; PID:9264861

R:Lee, B.; Godfrey, M.; Vitale, E.; Horl, H.; Mattei, M.G.; Sarfarazi, M.; Tsipouras,
 Nature 352, 330-334, 1991

A:Title: Linkage of Marfan syndrome and a phenotypically related disorder to two diff
 A:Reference number: S17062; MUID:91304567; PMID:1852206

A:Accession: S17062

A:Molecule type: mRNA

A:Residues: 'VLTVPVFIPLSYNKM', 944-1444 <LEE1>

A:Cross-references: EMBL:X62008; NID:931398; PIDN:CAB56534.1; PID:95924015

A:Accession: S62111

A:Molecule type: protein

A:Residues: 1166-1176, 'X', 1178-1180, 'D', 1182-1185 <LEE2>

R:Maddox, B.K.; Sakai, L.Y.; Keene, D.R.; Glanville, R.W.
 J. Biol. Chem. 264, 21381-21385, 1989

A:Title: Connective tissue microfibrils. Isolation and characterization of three larg
 A:Reference number: A34198; MUID:90078246; PMID:2512293

A:Accession: A34198

A:Molecule type: protein

A:Residues: 565-575; 1890-1892, 'T', 1894-1900 <MAD>

C:Comment: Fibrillin is a major component of elastin-associated microfibrils.

C:Gene: GDB:FBN1

A:Cross-references: GDB:127115; OMIM:134797; OMIM:154700

A:Map position: 15q21.1-15q21.1

A:Antons: 2236/1; 2258/1; 2297/1

C:Superfamily: fibrillin 1; EGF homology

C:Keywords: alternative splicing; calcium binding; extracellular matrix; glycoprotein
 F:1-3002/Product: fibrillin (5'-region exon A splice form) (fragment) #status predict
 F:1332-1367/Domain: EGF homology <EGF>

F:1457-1492/Domain: EGF homology <EGF2>

F:2262-2295/Domain: EGF homology <EGF1>

Query Match 21.7%; Score 523.5; DB 2; Length 3002;
 Best Local Similarity 30.3%; Pred. No. 9,5e-27;

Matches 122; Conservative 47; Mismatches 120; Indels 113; Gaps 18;

QY 1 OCTNGFDLDRSGGCLDDECRITPACRGDMCVNONGYLCIPRTNPNVNGPSNPYS 60
 DB 1185 RCDSEFALDSEBRNCTDIDECRISPLC-GRGQCVNTPGDFEC--KCDEGYESGF----- 1236
 QY 61 TPYSGPYAAPPLSAPNYPTISR-----PLICRF-----GYOMES-NO 99
 DB 1237 -----MMKNCMDIDECORPDLICGVCHNTEGYSRCECPGHOLSPNISA 1283
 QY 100 CVDVDECATDSHQCNPTQICINTEGGYTCSDGYMLEGQ--CLDIDECRY--GYCOOL 155
 DB 1284 CIDINECELSAHC-PNGRCVNLIGKYQACANNGYSTPRLCVDIDECISNNGGETF 1342
 QY 156 CANVGSYSCTGNPFTLNEDEGSCDVNVECATENP----- 191
 DB 1343 CTNSGSEYECSCQPGFALMPDRSCDIDDEC-EDNNICDGGQCTNIPGEYRCICYDFM 1401
 QY 192 -----CVQ-----TCVNTYGSFICRDPGELEEDGHCSDMDECSFSE 230
 DB 1402 ASEDKRTVDVNECLNPTICISGTCENTKGSFICDMDGYSKKKGTGTDINECEIGA 1461
 QY 231 FLC-QHE-CVNOPGTGFCSPGYILLDNRSQCDINECEHRNHTCNLOQTCYNLQGGFK 288
 DB 1462 HNCGHAVCTNAGSFKSCSPGWT--GGINCTDIDECSSNGTHMCSQAHDCNTMGSYR 1519
 QY 289 CIDPICEEPTL-----RISDNRCMPAENPCRDOP 320
 DB 1520 CL-----CREGYTGDFCTDIDECSENILIC--GNGQCLNAP 1555

submitted to the EMBL Data Library, June 1998

A:Description: Identification of chicken and C. elegans fibulin-1 homologs and character

A:Reference number: T22337

A:Accession: T43210

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-569 <BAR>

A:Cross-references: EMBL:AF070477; PIDN:AAC24035.1

C:Genetics:

A:Note: Intron positions not resolved (incomplete sequence)

C:Superfamily: unassigned EGF-related proteins; EGF homology

Query Match 24.6%; Score 592.5; DB 2; Length 589;

Best Local Similarity 27.4%; Pred. No. 6.3e-32;

Matches 156; Conservative 52; Mismatches 155; Indels 207; Gaps 20;

2 CTNGFDLRSGGCLD-IDECRTIPEACRGDMCMVNGGILCIPRTNPTVYRGPISNYS 60

79 CRSGFDLPAGMACVCHIDECATLMDCLSSQRCINPPGSPKCI----- 122

61 TPYSGYPAAPLAPLAPNPTISRLICRFGYOMD-ESNOCVDVDECATDSHOCNPTQIC 119

123 -----RTLSGIGIYAMDSETERCRDVEDCMGSHDCGPLYQC 159

120 INTEGTYC-----SCTDGYMLLEGQCLDIDECRYG----- 151

160 RMTGSGYRCDAKKGDELQNPMTGECTSTICPNGYVPKNGCMNDIDECVGHNGAGBE 219

152 -----CGQ---LCA-----NVPGSYCTCNPG 170

220 CVNTPGSPFRQKGNLCAMHGYVNGATGCEDEVNCOQVCGSMECIMLPGYKCKGSPG 279

171 FTLE-----DGRSCODVNEC 186

280 YENDAKKRCEDVEDCKIKFAGHYCDLSAECINTIGSEPCCKKPGFQLASDGRKEDVNEC 339

187 AFE-NPCVOTCVNTYSGFICRDPGYLEEDGVHCSMDDECSF-----SEFLQHECVNOP 241

340 TGIACCEGKCNVMPGSGYQCIDRGFALGPDGKCEDIDECSIMAGSGNDLGMGCIWTK 399

242 GTFYSGCPRYIILDDNRSCODINECEHNNHNTCMLOQTCYNLOGGFKCIDPICEPIYR 301

400 GSTLCCCPGPKIQPDGRVCVDVDECA-MGECAGSDKVCVNTLGSFK-HSIDCPNTYIH 457

302 ISDNR-----C-MCPAENPGC-RDOPFTIYRDMDVVSGRSV--PADI----- 340

458 DSLNKGQIADGSGICIVCSFEDIECLGNHTRVLYKQRAVPSLKTIIISPIEVSRIYTHMG 517

341 --FOMQATRYPGAYYIFQIKSGNEGREFTWROTGPISATLVMTRPRIKGPRIQDLEMT 398

518 VPFSSVDYNDLYVQQRHRIYQERNIG-----IVQLVKPISGP-----TVEYI 559

399 TVNTVINFSGSYIR-----LRIYVSQYRF 423

560 KVNIMTKSRGTGLAENEALIEISVSKYF 589

DB

RESULT 12

T22793

hypothetical protein F56H1.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Aug-2002

C:Accession: T22793; T24489

R:Wilkinson, J.

submitted to the EMBL Data Library, January 1996

A:Reference number: Z19616

A:Accession: T22793

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-798 <MIL>

A:Cross-references: EMBL:Z68749; PIDN:CAA92362.1; GSPDB:GN00022; CESP:F56H1.1

A:Experimental source: clone F56H1

R:Lloyd, C.

submitted to the EMBL Data Library, December 1995

A:Reference number: Z19897

A:Accession: T24489

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-798 <MIL>

A:Cross-references: EMBL:Z68219; PIDN:CAA92483.1; GSPDB:GN00022; CESP:F56H1.1

A:Experimental source: clone T05A1

C:Genetics:

A:Gene: CESP:F56H1.1

A:Map position: 4

A:Introns: 14/1; 92/3; 144/1; 195/1; 281/1; 325/2; 371/2; 390/1; 437/1; 471/3; 498/3;

C:Superfamily: fibulin-1; EGF homology

Query Match 23.8%; Score 574; DB 2; Length 798;

Best Local Similarity 26.8%; Pred. No. 1.4e-30;

Matches 159; Conservative 49; Mismatches 172; Indels 214; Gaps 22;

2 CTNGFDLRSGGCLD-IDECRTIPEACRGDMCMVNGGILC-----IP---RTNPTV 51

180 CRSGFDLPAGMACVCHIDECATLMDCLSSQRCINPPGSPKCI----- 239

52 RGP---YSNPTSPYSGPIPAAP----- 73

240 NABRRMRDDEYSR--AGEYRASAQANTEFGCPMGMLFQHGHCVIDECATLMDCLSSQ 297

74 -LSAPNPTISRLICRFGYOMD-ESNOCV-----DVDECATDSHOC 113

298 CLNTPGSPFKCIRITLSCGTGYAMDSETERNNCFLLIINNTFNCKTFVEDVDECMGSHDC 357

114 NPTQICINTEGTYCS----- 129

358 GPLYCCNNTGSGYRCDAKKGDELQNPMTGEYIDECVTHNGCAGBEVCNTPGSPRCQ 417

130 -----CTDGYML--LEGQCLDIDECRYGCOOL-CANVGSYSCTCNPGFTLE----- 175

418 KGNLCAMHGYVNGATGCEDEVNCOQVCGSMECIMLPGYKCKGSPGFEYFADAKKRED 477

176 -----DGRSCODVNECATE-NPCVQYCV 197

478 VDECIKFAHYCDLSAECINTIGSEPCCKKPGFQLASDGRKEDVNECTTIGIAAEQKCV 537

198 NTYSGFICRDPGYLEEDGVHCSMDDECSF-----SEFLQHECVNOPGTYFCSPGYI 253

538 NIGSYQCIDRGFALGPDGKCEDIDECSIMAGSGNDLGMGCIWTKGSLYQCPGPK 597

254 LDDNRSCODINECEHNNHNTCMLOQTCYNLOGGFKCIDPICEPIYRISDNR----- 306

598 IOPDGRVCVDVDECA-MGECAGSDKVCVNTLGSFK-HSIDCPNTYIHDSLNKRNCRNOP 655

307 --CMCPAENPGCRDOPFTIYRDMDVVSGRSV-----ADI---FQM 343

656 SAGGLPBE---CSKVPLEFLTYQFISL--ARAVPSSHRPATILFKVSAPNMADTEVNPEL 710

344 QATRYPGAYYIFQIKSGNEGREFTWROTGP--ISATLVMTRPRIKGPRIQDLEMT 395

711 QKTTIYGADNVLPALIRAN-----FLQKGEKRSAAVTVTLRDSLDGPDQTVKQL 759

DB

RESULT 13

A35626

transforming growth factor beta-1-binding protein - human

C:Species: Homo sapiens (man)

C:Date: 21-Sep-1990 #sequence_revision 21-Sep-1990 #text_change 11-Jan-2000

C:Accession: A35626

R:Kanzaki, T.; Olofsson, A.; Moren, A.; Wernstedt, C.; Hellman, U.; Miyazono, K.; Cla

Cell 61, 1051-1061, 1990

A:Title: TGF-beta1 binding protein: a component of the large latent complex of TGF-be

A:Reference number: A35626; MUID:90275601; PMID:2350783

A:Accession: A35626

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-1394 <KAN>

OY 232 ICGHECVNPGTYFCSCP-PGYILLDDNRSCODINECHRNHTCNLOOTCYNLOGGFKCI 290
 Db 493 ICYRCINIDGSPSCSSGYSGLAPNGRNCODIDECVTGIGHNSINETCNIDGAFKCL 552
 OY 291 DPICRCEPYLR 301
 Db 553 -AECPEENYR 562

RESULT 9

T42760
 fibulin, splice form D precursor - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
 C:Accession: T42760
 R:Barth, J.L.; Argaves, K.M.; Roark, E.F.; Little, C.D.; Argaves, W.S.
 submitted to the EMBL Data Library, February 1998
 A:Description: Isolation of chicken and nematode fibulin-1 homologs and characterization
 A:Reference number: 222267
 A:Accession: T42760
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-689 <BAR>
 A:Cross-references: EMBL:AF051401; PIDN:AAC28321.1
 C:Genetics:
 A:Note: FBLN1
 C:Superfamily: unassigned EGF-related proteins; EGF homology

Query Match 25.1%; Score 603; DB 2; Length 689;
 Best Local Similarity 27.4%; Pred. No. 1.5e-32;
 Matches 156; Conservative 52; Mismatches 155; Indels 206; Gaps 19;

OY 2 CTNGFDLDNOSGCLDIDECRTIPEACRGDMCVNONGYLCTPRTPYRGPYSNPYST 61
 Db 180 CRSGFDLPADGMACVDIDECATLMDCLDSQRCINLPGSFKCI----- 222
 OY 62 PYSGPYPAAPPLSAPNYPITISRLICRFGYOMD-ESNOCVDVDECATDSHOCNPQICI 120
 Db 223 -----RTLSCGTGYAMDSETERCRDVECMGSHDCGPLYQCR 260
 OY 121 NTEGGYTC-----SCTDGYMLLEGQCLDIDECRYG----- 151
 Db 261 NTQGSYRCDAKKCGDGLQNPMTGECTSITCPNGYYPKNCMDNDIDECVTGHNCGAGEC 320
 OY 152 -----CGQ--LCA-----NVPSYSTCNPGE 171
 Db 321 VNTPGSFRQKGNLCAHGEYVNGATGFCEDVNECOGVCMSMECINLPETYYCKCGPGY 380
 OY 172 TLNE-----DGRSCODVNECA 187
 Db 381 EENDAKKREDEDECIFAGHVODLSAECINTIGSFECCKPFGOLASDGRCEDVNECT 440
 OY 188 TE-NPCVQTCVNTYGSFTRCDPGYLEEDGVHCSMDDECSF-----SEFLQHECVNQP 242
 Db 441 TGIACCEKCVNIPESYQICIDRGFALGPDGTKECEDIDECISIWAGSNDLCMGGCINTKG 500
 OY 243 TYFGSCPFGYILLDDNRSCODINECHRNHTCNLOOTCYNLOGGFKCIDIRCEEPYLR 302
 Db 501 SYLCCPFGYKIQPGRTCVDVDECA-MGECASDVKCVNTLSGFKC-HSIDCPNTYIHD 558
 OY 303 SDNR-----C-MCPAENPGC-RDQFTILYRDMVVSGRSV--PADI----- 340
 Db 559 SLNKNQIADGVSCIVCSTEDTECGNHNREVLYQRAVPSLTIISPIEVSRTVTHMGV 618
 OY 341 -FOMQATTRYPGAYITIFQIKSNEGREFYMRGTGPISATLWMPRIKGPRIOLDLEMIT 399
 Db 619 PESVDYNDYVGQRHFRIYQERNIG-----IVQIVKPIGSP-----TYETIK 660
 OY 400 VNTVINFRGSSVIR-----LRIYVSQYF 423
 Db 661 VNHTKSRGTGYLANEALITELTSKVPYF 689

RESULT 10

T42990
 fibulin 1, splice form C precursor - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
 C:Accession: T42990
 R:Barth, J.L.; Argaves, K.M.; Roark, E.F.; Little, C.D.; Argaves, W.S.
 submitted to the EMBL Data Library, February 1998
 A:Description: Isolation of chicken and nematode fibulin-1 homologs and characterization
 A:Reference number: 222267
 A:Accession: T42990
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-712 <BAR>
 A:Cross-references: EMBL:AF051402; PIDN:AAC28322.1
 C:Genetics:
 A:Gene: FBLN1
 C:Superfamily: unassigned EGF-related proteins; EGF homology
 C:Keywords: alternative splicing; basement membrane; extracellular matrix

Query Match 24.8%; Score 596; DB 2; Length 712;
 Best Local Similarity 27.9%; Pred. No. 4.4e-32;
 Matches 152; Conservative 46; Mismatches 145; Indels 202; Gaps 18;

OY 2 CTNGFDLDNOSGCLDIDECRTIPEACRGDMCVNONGYLCTPRTPYRGPYSNPYST 61
 Db 180 CRSGFDLPADGMACVDIDECATLMDCLDSQRCINLPGSFKCI----- 222
 OY 62 PYSGPYPAAPPLSAPNYPITISRLICRFGYOMD-ESNOCVDVDECATDSHOCNPQICI 120
 Db 223 -----RTLSCGTGYAMDSETERCRDVECMGSHDCGPLYQCR 260
 OY 121 NTEGGYTC-----SCTDGYMLLEGQCLDIDECRYG----- 151
 Db 261 NTQGSYRCDAKKCGDGLQNPMTGECTSITCPNGYYPKNCMDNDIDECVTGHNCGAGEC 320
 OY 152 -----CGQ--LCA-----NVPSYSTCNPGE 171
 Db 321 VNTPGSFRQKGNLCAHGEYVNGATGFCEDVNECOGVCMSMECINLPETYYCKCGPGY 380
 OY 172 TLNE-----DGRSCODVNECA 187
 Db 381 EENDAKKREDEDECIFAGHVODLSAECINTIGSFECCKPFGOLASDGRCEDVNECT 440
 OY 188 TE-NPCVQTCVNTYGSFTRCDPGYLEEDGVHCSMDDECSF-----SEFLQHECVNQP 242
 Db 441 TGIACCEKCVNIPESYQICIDRGFALGPDGTKECEDIDECISIWAGSNDLCMGGCINTKG 500
 OY 243 TYFGSCPFGYILLDDNRSCODINECHRNHTCNLOOTCYNLOGGFKCIDIRCEEPYLR 302
 Db 501 SYLCCPFGYKIQPGRTCVDVDECA-MGECASDVKCVNTLSGFKC-HSIDCPNTYIHD 558
 OY 303 SDNR-----CMCPAENPGCRDQFTILYRDMVVSGRSV----- 337
 Db 559 SLNKNRCARQPSACGLPEE--CSKVPFLTLTYQFISL--ARAVPISHRAITLFRVSAP 613
 OY 338 --ADI--FOMQATTRYPGAYITIFQIKSNEGREFYMRGTGP--ISATLWMPRIKGPRI 390
 Db 614 NHADDEVNFEQLTKTIVGAPVYLAIRAN-----FLQKGERKNSAVVTLRSLDGPQT 668
 OY 391 IOLD 395
 Db 669 VKQL 673

RESULT 11

T43210
 fibulin-1D precursor - Caenorhabditis elegans (fragment)
 C:Species: Caenorhabditis elegans
 C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 02-Sep-2000
 C:Accession: T43210
 R:Barth, J.L.; Argaves, K.M.; Roark, E.F.; Little, C.D.; Argaves, W.S.

Best Local Similarity	35.28;	Pred. No.	4.5e-39;
Matches	160;	Mismatches	154;
Conservative	66;	Indels	74;
		Gaps	21

```

QY      1  OCTGFGDLDRSGGCLDIDECETPEACRGDMKVCNONGCYLICIPRTNVTNVRGPRYNPSS 60
Db      295  OCKSGFIOD -ALGNCLDIETCLISTAPRCVGGOTCINTBESYTC----- 336
QY      61  TPYSGPYPAAPRLSAPNYPPTISRELLICRGFYQM -DESNOCVDVDECANDSHOCNPTOIC 119
Db      337  -----QKNVPR----- -GGRGYHMBEGTRCYVDVDECAPRAEPCKAGHHG 375
QY      120  INTEGTYCTCSTDGYWL -LEGQCLDIDEC -RY--GYCQQLANPAGSISCTCNPFETLN 174
Db      376  LNSGSEFCEKAGYFDGISRTVDINECOQYPERLCCHKEMKPPSHSCSSAFAFRIS 435
QY      175  EDGRSCODVINCATEMPCVQTQVNTNYSFICRCRDGELEE -DGVHCSMDDECSF--SEF 231
Db      436  VDRGSCEDVNEC -LNSPCSQECANYYGSGYQCYCRGQYQSLSDVDGYTCEDIDEALPTTGGH 494
QY      232  LCGHECVNQOPRTYRCSCP -PGYIILDDNRSCODINDECHEHRNTCNLQOTCYNLQGFKCI 290
Db      495  ICSYRCINIPSPSCSPSSGSRRLAPNCRNRCODIDECVYGINCSINTECFNIOGFSRCL 554
QY      291  DPICEEPEYLRISDN-----RCM--CPAENPGC -RDQFTILYRMDVYSGRSV-- 336
Db      555  S-FECPENYRRSADTFREKTDVYACIKSCRRNDACVADPVNHYSHYVYISLTFEFTR 613
QY      337  PADIFQOQAT -RYPG -AAYIFQIKSGNEGHEFYM--RQTPISATLWMPRIKGP 390
Db      614  PEETIFLAAYPLPAPNQAQDIIFDTTEGNLRDSFDIKRYEGDMGTGVAVRVPIGPFY 673
QY      391  IQLLLEM -IYNTYINFRGSSVIRLRIYVSOQPF 423
Db      674  AVLTLEMYVLGGVYSHR -NVNVNHYIEFSEWF 705

```

RESULT 7
A:55184
fibulin-2 precursor - human
N:Alternate names: protein DKFZP586A1519.1
C:Species: Homo sapiens (man)
C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 02-Aug-2002
C:Accession: A55184; S087184
R:Zhang, R.Z.; Pan, T.C.; Zhang, Z.Y.; Mattei, M.G.; Timpl, R.; Chu, M.L.
Genomics 22, 425-430, 1994
A:Title: Fibulin-2 (FBLN2): human cDNA sequence, mRNA expression, and mapping of the gene
A:Reference number: A55184; MUID:95104855; PMID:7806230
A:Accession: A55184
Status: preliminary
Molecule type: mRNA
M:Residues: 1-1184 <ZNA>
A:Cross-references: GB:X82494; NID:g575232; PIDN:CAA57876.1; PID:g575233
R:Wambutt, R.; Haubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, March 1999
A:Reference number: Z16471
A:Accession: T08744
A:Molecule type: mRNA
A:Residues: 655-719, <ODECLMGADHCNRQECVNTLGSEFYCVNHTVLCADGTYILNHRKQVD', 720-853, 'T', 855-1184>
A:Cross-references: EMBL:AL050095
A:Experimental source: adult uterus; clone DKFZP586A1519
C:Genetics:
A:Gene: GDB:FBLN2
A:Cross-references: GDB:293037; OMIM:135621
A:Map position: 3p25-3p24
A:Note: DKFZP586A1519.1
C:Superfamily: fibulin-2; EGF homology
C:Keywords: alternative splicing; extracellular matrix
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-1184/Product: fibulin-2 protein #status predicted <MAT>
F:905-941/Domain: EGF homology <EGF>

Matches 142; Conservative 58; Mismatches 152; Indels 59; Gaps 11;

```

QY 1 QCNTEPFLDQSGGCLIDECRTIPKACRGDMMCVNONGYLCTPRNPVYRGYSNPS 60
Db 795 RCMDF-LQDPBGKCVINDICTSISEPCRGFGSCINVGSTYC----- 836
QY 61 TPYSGPYAPAAPLSAPNYPITISRPLICRFYQ-MDESNOCVDVDECATSDHOCNPJOIC 119
Db 837 -----QNRPLICARHYHSSDDGAKVDVNECEYGVHRCGEQVC 875
QY 120 INTEGGTSCSTDYW--LLEGCLDIDECRYG--YCQOLCANVPGSYSTCNPFGFLN 174
Db 876 HNLFSGYRCCKAKAFQDADAFGRGCI DVNECWASPGRLCQHNCEVTLSTYRSCASGFLTA 935
QY 175 EDGRSCQDVNECATENPCVOYICVVTYSFICRCRCPGYELEDEYGVCSMDCCS-FSEPIC 233
Db 936 ADGRCEVNECEMQR-CSQECANITYSYOCTCYOGYQJALADHGHTCTDIDBCAGAGATLIC 994
QY 234 QHECVNPGYFCSCP-PGYILLDDNRSQDINECEHRNHTCNLQOTCYNLQGGFKIDP 292
Db 995 TFRCLANPGSYQACPRGYMTMTANGRSCKVDSCALGTNHCSEAKETHNIOGSFRCL-R 1053
QY 293 IRCEPPLRLISDNRCMKPAENPCGRD-----QPTLLYKMDVYSGRSVPADLFOQAT 346
Db 1054 FECPPNVOVSKTCC-----ERTTCHDELCEONSPARITHYOLNQTGLVPAHFLRIGPA 1109
QY 347 TRYRGATVYIPKISGNEGREFYMKOTGPISATLVMTPIRKIPREIOLDLEM 397
Db 1110 PAFEGDTIALNLKGNBEGYGTIRRLNAYTGVTLQRAVLEPRFADLVEM 1160

```

RESULT 8
B36346
fibulin 1 precursor, splice form B - human
C:Species: Homo sapiens (man)
C:Date: 19-Apr-1991 #sequence_revision 19-Apr-1991 #text_change 02-Aug-2002
C:Accession: B36346
R:Argaves, W.S.; Tran, H.; Burgess, W.H.; Dickerson, K.
J. Cell Biol. 111, 3155-3164, 1990
A:Title: Fibulin is an extracellular matrix and plasma glycoprotein with repeated dom
A:Reference number: A36346; M01D:91100426; PMID:2268665
A:Accession: B36346
A:Molecule type: mRNA
A:Residues: 1-601 <ARG>
A:Cross-references: GB:X53742; NID:g31416; PIDN:CA37771.1; PID:g31417
C:Genetics:
A:Gene: GDB:FBLN1; FBLN
A:Cross-references: GDB:278285; OMIM:135820
A:Map position: 22q13.3-22q13.3
C:Superfamily: fibulin-1; EGF homology
C:Keywords: alternative splicing
F:180-214/Domain: EGF homology <EGF1>
F:485-523/Domain: EGF homology <EGF>

	Query Match	25.7%	Score 617.5	DB 2	length 601
	Similarity	38.3%	Pred. 1.4e-33		
	Matches	119	Conservative	46	Mismatches 95; Indels 51; Gaps 12
QY	1	QCTNGFDIDROSGGOCIDIDECRTTIEACRGMDMCVNQNGYLICIPPTNPYRGPSNPYS	60		
DB	293	QCKSGFIQD-ALGNCIDINECLISAPCPDIGHICNTGESSYC-----	334		
QY	61	TPYSGPYAAAPPLPSAPNPPTISRLPICFQGM-DESMQCVYDVDECAIDSHQCNPTQIC	119		
DB	335	-----QKNVNP-----GGRGHLINEEGRVDVDECAAPAEPCGGHRC	373		
QY	120	INTEGGYVNSCTDGYWL--LEGQCLDIDEC-RY--GYCOOLCANVSGSYCTCNPGFTLN	174		
DB	374	VNSPGRFCECKTGYTFDGIISMCMVDVNECQYRGRRLCGHKCENTLGTSLCSCSYGPRLS	433		
QY	175	EDGSSQCVNFCATINPCVQTCVNTVNGSTICGCDPQYELEF--DGVNCSMDDESF--SEF	231		
DB	434	VDGNSCEIDINCS--SPCSQECANVAGSYQCYCRKRYTQSDVDGYVCEIDICALPTGGH	492		

A:Residues: 30-35, 'SX', 38-40, 'SH', 43-44 <AR3>
 A:Genetics:
 A:Gene: GDB:FBLN1; FBLN
 A:Cross-references: GDB:278285; OMIM:135820
 A:Map position: 22q13.3-22q13.3
 C:Superfamily: fibulin-1; EGF homology
 C:Keywords: alternative splicing; glycoprotein
 F:1-29/Domain: signal sequence status predicted <SIG>
 F:30-683/Product: fibulin 1 splice form C #status predicted <MAP>
 F:180-214/Domain: EGF homology <EGF>
 F:485-523/Domain: EGF homology <EGF>
 F:98,535,539/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 30.6%; Score 735.5; DB 2; Length 683;
 Best Local Similarity 35.7%; Pred. No. 2,7e-41;
 Matches 155; Conservative 66; Mismatches 152; Indels 61; Gaps 16;

```

QY 1 OCTNGFPLDROSQCCLDIDECRTTPEACRGDMCVNONGYLCTPRTNRYRGPSNYS 60
DB 293 QCKSGFID-ALGNCLIDNECLISAPCPIGHHTCINTGSSYTC----- 334
QY 61 TPYSGPYPAAPPLSAPNYPITISRLICRGYOM-DESNOCVDVDECATDSHCNPTQIC 119
DB 335 -----QKNVPR-----CGRGTHNEEGTRCVDECAPAEPCGKGRHC 373
QY 120 INTEGYTSCCTDGYW-LEGGCLDIDEC-RY-GYCOQLCANVPSYSTCNPFTLN 174
DB 374 VNSGSRCECKTGTYFDGISRMCDVNEGCRGRLGHCNENTLGSITLCSGIFRLS 433
QY 175 EDGSCODVNECATENPVQTCVNTYGSFICRDPGELEB-DGVHCSMDCESE--SEF 231
DB 434 VDRGSCDIDNECS-SPCSQCANVGSYOCYCRGQQLSDVDGTCEDIDECALPTEGH 492
QY 232 LCHCEVNOBQTYFCSCP-PGYILLDNRSQDINECEHRNHTNLTQTCYNIGSGFCI 290
DB 493 ICYRCININGSPQCSGSSGYRLAPNGRNCQDIDECVTGHNCSINETCNIGQAEFCL 552
QY 291 DPICRCEPYLRISDNRC-MCPA-ENPGCRDPTILYRDMVVSGRSVPADIFOMQATTR 348
DB 553 -AECEENRRSATRERLRPECHENRECSKPLRITYYHSFPNIOAPAVFVRMGPSA 611
QY 349 YGAYVYIFQIKSGNEGREFYRQTPISATLVMRPKIPGKEIOLDLMTVNVINRG 408
DB 612 VPDSNQLAITGNNEGEFFTRKRVSPHSGVALTKVPEDL-----LTVKMDLSNHG 666
QY 409 -----SSYIRLRIYVS 419
DB 667 TVSSFVAKLFIYVS 680

```

RESULT 5

A:9457
 A:Residues: 1-1221 <PAN>
 A:Molecule type: mRNA
 A:Cross-references: GB:J75285; NTD:9437046; PIDN:CAAS3040.1; PID:9437047
 R:Sasaki, T.; Mann, K.; Murphy, G.; Chu, M.L.; Timpl, R.
 Eur. J. Biochem. 240, 427-434, 1996
 A:Title: Different susceptibilities of fibulin-1 and fibulin-2 to cleavage by matrix met
 A:Reference number: S74094; MUID:96439073; PMID:8841408
 A:Accession: S74095
 A:Molecule type: protein
 A:Residues: 236-238, 'X', 240-247, 260-275, 336-344, 'L', 346-361, 405-426, 566-568, 'EW', 569-589
 C:Superfamily: fibulin-2; EGF homology

C:Keywords: calcium binding; duplication; extracellular matrix; glycoprotein; homotr
 F:942-978/Domain: EGF homology <EGF>

Query Match 30.3%; Score 729.5; DB 2; Length 1221;
 Best Local Similarity 35.3%; Pred. No. 1.1e-40;
 Matches 145; Conservative 59; Mismatches 148; Indels 59; Gaps 11;

```

QY 1 OCTNGFPLDROSQCCLDIDECRTTPEACRGDMCVNONGYLCTPRTNRYRGPSNYS 60
DB 832 RCMDFG-LDPREGVCVINECTSLIEFCRSFCSCINTVGSTTC----- 873
QY 61 TPYSGPYPAAPPLSAPNYPITISRLICRGYOM-DESNOCVDVDECATDSHCNPTQIC 119
DB 874 -----QRNPLVCGRGYHANESECVNDECTGVHRCGEGLC 912
QY 120 INTEGYTSCCTDGYW-LEGGCLDIDECRYG---YCQQLCANVPSYSTCNPFTLN 174
DB 913 YNLPSTYRCDCRKGFORDAFGRTCIDNECVWSPGRICQHTCENTPSSYRSCAAGFLA 972
QY 175 EDGSCODVNECATENPVQTCVNTYGSFICRDPGELEBDEGVHCSMDCESE-FSEFIC 233
DB 973 ADGKCEDEVNECESTR-CSQECANITGSGYOCRGQQLADHDHTCTDIDECACAGACILC 1031
QY 234 QHCEVNOBQTYFCSCP-PGYILLDNRSQDINECEHRNHTNLTQTCYNIGSGFKCIDP 292
DB 1032 TFRCVNVPSSYOCACEQGTMMANRSCKDIDECALGTNCSAEATCNIGSEFRL-R 1090
QY 293 ICEPEYLRISDNRCMPAENPGCRD-----OPFTLYRDMVVSGRSVPADIFOMQAT 346
DB 1091 FDCPPRYVAVSQTKC-----ERTTCQDITECQTSPTATHTQLNFGGLPAHIFRIGPA 1146
QY 347 TPYGAIVYIFQIKSGNEGREFYRQTPISATLVMRPKIPGKEIOLDLMTVNVINRG 397
DB 1147 PAFADDTISLTITKNGEEGYFTRRLNATGVSLORSYLEPRDFALDYEM 1197

```

RESULT 6

A:9468
 A:Residues: 1-705 <PAN>
 A:Molecule type: mRNA
 A:Cross-references: EMBL:X70854; NTD:9396820; PIDN:CAAS0207.1; PID:9396821
 R:Kluge, M.; Mann, K.; Dziadek, M.; Timpl, R.
 Eur. J. Biochem. 215, 733-740, 1993
 A:Title: Sequence of extracellular matrix protein BM-90/fibulin and its calcium-depend
 A:Reference number: S34968; MUID:93358897; PMID:8354280
 A:Accession: S34968
 A:Molecule type: protein
 A:Residues: 1-705 <PAN>
 R:Pan, T.C.; Kluge, M.; Zhang, R.Z.; Mayer, U.; Timpl, R.; Chu, M.L.
 submitted to the EMBL Data Library, January 1993
 A:Description: Sequence of extracellular matrix protein BM-90/fibulin and its calcium-
 A:Reference number: S36440
 A:Accession: S36441
 A:Molecule type: mRNA
 A:Residues: 1-39, 'P', 41-705 <PAN>
 A:Cross-references: EMBL:X70854; NTD:9396820; PIDN:CAAS0207.1; PID:9396821
 R:Kluge, M.; Mann, K.; Dziadek, M.; Timpl, R.
 Eur. J. Biochem. 193, 651-659, 1990
 A:Title: Characterization of a novel calcium-binding 90-kDa glycoprotein (BM-90) shar
 A:Reference number: S13814; MUID:91065369; PMID:2249686
 A:Accession: S13814
 A:Molecule type: protein
 A:Residues: 28, 31-49, 'X', 51-53, 'XX', 110-117, 231-240, 'X', 242-243, 339-362, 'S', 364-387, 4
 C:Superfamily: fibulin-1; EGF homology
 C:Keywords: alternative splicing; basement membrane; calcium binding; extracellular m
 F:1-29/Domain: signal sequence #status predicted <SIG>
 F:30-705/Product: fibulin, splice form D #status predicted <MAP>
 F:98,537,541/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 29.2%; Score 702; DB 2; Length 705;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 17:36:12 ; Search time 16.9977 Seconds
(without alignments) 2392.374 Million cell updates/sec

Title: US-09-674-379A-14

Perfect score: 2407

Sequence: 1 QCTNGFDLDRSGCCLDIDE.....INFRSSVIRLRIVSYQYF 423

Scoring table: BLASTOM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1108.5	46.1	493	2 JC5621	epidermal growth f
2	963.5	40.0	387	2 I38449	extracellular prot
3	744.5	30.9	685	2 S78040	fibulin, splice fo
4	735.5	30.6	683	2 C36346	fibulin-1 precursor
5	729.5	30.3	1221	2 A49457	fibulin-2 precursor
6	702	29.2	705	2 S34968	fibulin, splice fo
7	701.5	29.1	1184	2 A55184	fibulin-2 precursor
8	617.5	25.7	601	2 B36346	fibulin-1 precursor
9	603	25.1	689	2 T42760	fibulin, splice fo
10	596	24.8	712	2 T42990	fibulin-1, splice
11	592.5	24.6	589	2 T43210	fibulin-1D precurs
12	574	23.8	798	2 T22793	hypothetical prote
13	539	22.4	1394	2 A35626	transforming growt
14	523.5	21.7	3002	2 A47221	fibillin-1 precur
15	522.5	21.7	2871	2 A55567	fibillin-1 precu
16	518	21.5	1712	2 A55624	masking protein pr
17	509.5	21.2	2871	2 A55624	fibillin-1 precu
18	509.5	21.2	2907	2 A55624	fibillin-2 precu
19	508.5	21.1	2918	2 A54105	latent transformin
20	469.5	19.5	1820	2 A54944	hypothetical prote
21	455	18.9	741	2 T46488	hypothetical prote
22	448.5	18.6	1620	2 T27283	hypothetical prote
23	438	18.2	1251	2 A57293	latent transformin
24	436	18.1	1574	2 T13954	MEGF6 protein - ra
25	392	16.3	886	2 A57172	probable hormone r
26	384.5	16.0	3507	2 T34513	hypothetical prote
27	368.5	15.3	1106	2 T18739	hypothetical prote
28	367	15.2	2471	2 A49128	cell-fate determin
29	364	15.1	810	2 T10756	Nel-homolog protei

30	361	15.0	1081	2 T31329	receptor tyrosine
31	359	14.9	1203	2 A49175	Notch B protein -
32	356	14.8	2552	2 A40043	Notch protein homo
33	353.5	14.7	1964	2 T09059	Notch4 - mouse
34	351.5	14.6	2437	2 S42612	transmembrane prot
35	348.5	14.5	2703	1 A24420	Notch protein - fr
36	340.5	14.1	673	2 A48089	growth arrest-spec
37	340	14.1	2524	2 A35844	Notch protein - Af
38	335.5	13.9	2531	2 S18188	Notch protein homo
39	335	13.9	674	2 T17298	growth potentialin
40	334	13.9	511	2 T17298	hypothetical prote
41	334	13.9	2531	2 T31070	Notch homolog - se
42	333	13.8	1064	2 A40136	fibropellin Ia - s
43	333	13.8	2381	2 S78549	Notch3 protein - h
44	325.5	13.5	2531	2 A46019	Notch-1 protein -
45	324.5	13.5	2318	2 S45306	Notch 3 protein -

ALIGNMENTS

RESULT 1

JC5621.

epidermal growth factor-like protein, r16 precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 09-Oct-1997 #Sequence_Revision 07-Nov-1997 #Text_Change 05-Nov-1999

C:Accession: JC5621

R:Ozaki, T.; Kondo, K.; Nakamura, Y.; Ichimura, S.; Nakagawara, A.; Sakiyama, S.

Biochem. Biophys. Res. Commun. 237, 245-250, 1997

A:Title: Interaction of DA41, a DAN-binding protein, with the epidermal growth factor

A:Reference number: JC5621; MUID:97415782; PMID:9268694

A:Accession: JC5621

A:Molecule type: mRNA

A:Residues: 1-493 <OZA>

A:Cross-References: DBJ:089730; NID:q2429082; PIDN:BAA22265.1; PID:di023127; PID:q24

C:Comment: This protein plays a role in the regulation of cell growth by interacting

C:Keywords: glycoprotein

F:1-17/Domain: signal sequence #status predicted <SIG>

F:128-70/Domain: signal sequence #status predicted <SIG>

F:249/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	Score	Length	DB 2	Length	DB 1
Best Local Similarity	46.1%	493	43.6%	56-66	5
Matches	204	Conservative	68	Mismatches	149
				Indels	47
				Gaps	5
QY	1	QCTNGFDLDRSGCCLDIDE	CRTPACRGMVQNGVLCIPRTNPNVGRPSNPS	60	
DB	28	QCTGDEWDPVROCKDIDEDIV	DACKGKCVNHVGYCLPRTAQIIVNNEPQOE	87	
QY	61	TPYS-----	GPYAAAPPLSAPNYP-----	81	
DB	88	TPAABASSGAATGTTAANSMT	SGVTPGGGFTASTAAVAGPVGGRNNFVIRRPADPO	147	
QY	82	-----ISRPILCRFGYQMD	ESNOCVDECATDSHCNPTQICITMTEGGYCTGTGYL	136	
DB	148	RIPSNPSHRICAGAGYEGE	SEHNVCDDICSGTINCRDVCINLRGFTCHCLPGYK	207	
QY	137	LEGCCLDIDECRY-GLY	COQLCANVPGYSTCNCNPEFTLNEDGRSCQDVNECATENPCVOT	195	
DB	208	RGECCVVIDECSVPYCHQ	GCVMTPGSFYCCCNPEFOIANNYTCDVIDNECDASNOCAQO	267	
QY	196	CVNTYGSFICRDEGL	DEEDGVCSDMECSFELCOHECVNOBGRYFGCSPGYTLL	255	
DB	268	CYNLTGSFICRDEGL	DEEDGVCSDMECSFELCOHECVNOBGRYFGCSPGYTLL	255	
QY	256	DDNRSCDINECEHRN	HTCNLQOTCYNLQGGFKCIDPRCEPYLRISDNRCMCPAENPG	315	
DB	328	-RSRTCODINECEHT	NE-CEDEMCWYHGGFRCPQPNQCDPYLTSENNCVCVSTNM	385	
QY	316	CRDPPFTILYRDM	YVGRSPADIFQMAITRRPGATYITQISGNGREFYMQTPPI	375	
DB	386	CRDPPFTILYRDM	YVGRSPADIFQMAITRRPGATYITQISGNGREFYMQTPPI	375	

FEATURES Tel:81-92-642-5466, Fax:81-92-642-5482)

Location/Qualifiers

1..1536

source /organism="Homo sapiens"

gene /db_xref="taxon:9606"

1..1536

/gene="MBP1"

97..1428

/gene="MBP1"

/note="Human homologue of murine mutant p53 binding protein 1 (MBP1) reported by Gallagher et al. (Oncogene 18: 3608-16, 1999)"

/product="mutant p53 binding protein 1 (MBP1)"

/protein_id="BA92880.1"

/db_xref="GI:7328921"

/translation="MLPCTSLPGLSLLMLLILLSASPODSEPDSTYECTDYE WPDSDHCRDNBCLTIPEACKGEMKINHGYCLIPRSAVINDLGEPPVPV AOHNPCEPPEPDSDSDVDCECAQALHDCRSODHNLPGSYOCTCPGPKRIG ECVDIECRARYCOHRCVNLPGSPRCOCEPFOGLPNNRSCVDNEDMDMAGPCBRCF NSYGFRCRHOGYELHRDGFSCSDIDECSSYLCOYRCVNEBREFCHCPQIQL ATRLCODIDCESGAHCSRAQTCVNFHGRVCRVDTNRCVPEYIIOVSENKCLCPASNP LCHEQPSYIHRMTITSERSVPADVFQIDTSVPAVNAFQIRAGNSQDFYIROI NNVSAMLVILARPVTGPREYVLDLEMTNLSMYSRASSVILRLTFVGAYTE"

BASE COUNT 307 a 500 c 436 g 293 t

ORIGIN

Query Match 29.4% Score 372.8; DB:9; Length 1536;

Best Local Similarity 61.3% Pred. No. 8.5e-96;

Matches 618; Conservative 0; Mismatches 387; Indels 3; Gaps 1;

262 TGCCGCTTGGATACGATGATGAAGCAACCAATGTGTGATGAGACGATGCA 321
 421 TGCCCAACAGGCTATGAGCCGACAGATCAGACAGCTGTGTGATGTGACGATGCC 480
 322 ACAGATTCCACAGTGCACACCCACCCAGATTCGATCACTACTGAAAGCGGTACAC 381
 481 CAGGCGCTGCACACGCTGCGCCAGCAGAGCTGCATTAAGTCCGCTCCATACAG 540
 382 TGCTCTGACGAGGATTTGGCTTCTGGAAGCCAGTGTAGATGATGATGT 441
 541 TGACCTGCTGATGTATGATGATGATGATGATGATGATGATGATGATGATGATG 600
 442 CGCTATGTTTACGACGAGCTGTGTGCAATGTTCTGATCTTCTGTATGATGC 501
 601 CGCTACCGCTACGACGAGCAGCGTGTGATGATGATGATGATGATGATGATGATG 660
 502 AACCTGGTTTACCTCAATGAGATGAGAGTCTTGCACAGATGTAAGAGATGTC 561
 661 GAGCGGGGTTTCCAGTGGGCTTAACACCGCTCTGTGTGATGTAAGAGATGTAAC 720
 562 ACCGAGAACCCCTGCTGCAACACCTGCTGCAACACCTGCTGCTGCTGCTGCTGCT 621
 721 ATGGGGGCCCCATGACGACGAGCTGCTTCACTGATGAGACCTTCTGTGTGCTGC 780
 622 GACCCAGATTAAGACTTGAAGATGAGGCTTCACTGATGATGATGATGATGATGATG 681
 781 CACGAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
 682 TTCTGTGATGCTGCTGCAACATGATGATGATGATGATGATGATGATGATGATGAT 741
 841 TACTCAGCTACTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
 742 TGCCCTCAGGCTATCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 801
 901 TGCCCAACAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 957
 802 GAGCAGAGAACACAGCTGACAGCTGACAGAGCTGCTACATTTTCAAGGGGCTTC 861
 958 GAGTCTGTGTCGACAGCTGTCGAGGCGCAAACTGTGTCAACTTCCATGGGGCTAC 1017
 862 AATGTCATGACCCCATCGCTGTGTGAGAGCTTATCTGAGATCACTGATTAACCGCTG 921

1018 CGCTGCTGACACCAACCGCTGCTGAGAGCCCTATATCCAGGTCGTGAGAACCGCTGT 1077
 922 ATGTGCTGCTGAGAACCCCTGCTGAGAGACCCCTTACCATCTTGTATACGGGAC 981
 1078 CTCTGCGCGGCTCCCAACCCCTGTGTGTGAGAGAGAGCCCTTATCATCTTGTGACCGCTAC 1137
 982 ATGAGCTGTGTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1041
 1138 ATGACCATCACTGCGAGCGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1197
 1042 CGTACCTGCGGCTTATCATCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1101
 1198 GTTACCTGCGGCTTATCATCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1257
 1102 TACATGCGGCAACGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1161
 1258 TACATGAGCAATCAACCAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1317
 1162 CCGCGGGAATCACTGAG 1221
 1318 CCGCGGAGATGAG 1377
 1222 GCGAGCTGCTGATGAG 1269
 1378 GCGAGCTGCTGATGAG 1425

Search completed: July 3, 2003, 15:34:28
 Job time : 3310.01 secs

ACCESSION AJ132819
 VERSION AJ132819.2 GI:4456472
 KEYWORDS FBLN4 gene; fibulin-4.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1440)
 AUTHORS Glatay, R., Timml, R. and Kostka, G.
 TITLE Sequence, recombinant expression and tissue localization of two novel extracellular matrix proteins, fibulin-3 and fibulin-4
 JOURNAL Matrix Biol. 18 (5), 469-480 (1999)
 MEDLINE 20068041
 PUBLISHED 10601734
 REFERENCE 2 (bases 1 to 1440)
 AUTHORS Kostka, G.
 TITLE Direct Submission
 SUBMITTED (11-FEB-1999) Kostka G., Proteinchemie,
 Max-Planck-Institut fuer Biochemie, Am Klopferspitz, 82152
 Martinsried, FRG
 COMMENT On Mar 21, 1999 this sequence version replaced gi:4454234.
 SOURCE Location/Qualifiers
 1..1440
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="melanoma"
 1..1332
 /gene="FBLN4"
 1..1332
 /gene="FBLN4"
 /codon_start=1
 /product="fibulin-4"
 /protein_id="CA10791.2"
 /db_xref="GI:4456473"
 /db_xref="SPTREMBL:O95967"
 /translation="MLPCTSLPESLLMALLLLSASQDSEEDPSYECTDYE
 WPDSDGCRDVNCLTIPKCKGEMKINHYGLCLPRAVAINDLGEPPPPVP
 AOHNPCEPGEYEDDDSCVDVDEACALHDCRPSODCHNLPSCYOCQPGYRKGIG
 EGVDDIECRHYRCDQHCNVNLPSCGCEGFCGLPNRSCVDNECDMGAEQRCF
 NSEYGLRCRHHQGLHEDGSCSDIDECSSYLCOYRCVNEPGRFCHPOGYOL
 ATRICDDIECESGAGHOCSEAFORCVNHHGGRVCYDTRCVPEYIOVSENRCLCPASNP
 LCRBPPSSIVHRTMTTSESVADYFQIQTATSPRYANAFQIRAGSGDFITROI
 NNVSAVLARPTGPREYVLEMTVMSLSMRASSVLRLTFVGAATP"
 1..81
 /gene="FBLN4"
 82..1329
 /gene="FBLN4"
 /product="fibulin-4"
 BASE COUNT 291 a 467 c 397 g 285 t
 ORIGIN
 Query Match 29.4%; Score 372.8; DB 9; Length 1440;
 Best Local Similarity 61.3%; Pred. No. 8.5e-96;
 Matches 618; Conservative 0; Mismatches 387; Indels 3; Gaps 1;

Db 565 GACCGGGTTTCCAGCTGGGGCTTAACACCGCTCTGTGTGATGTAACGAGTGTGAC 624
 QY 562 ACCGAGAACCCCGCGCTGCAAAACCTGGCTTAACACCTGGCTTTGATCGCGGCT 621
 Db 625 ATGGGGGGCCCATGCGAGCAGCGCTGCTTCACTCACTTATGAGACCTGTGTGCTG 684
 QY 622 GACCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 681
 Db 685 CACAGGGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 744
 QY 682 TTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 741
 Db 745 TACTTCAAGTACCTCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 804
 QY 742 TGCCCTCCAGGCTTACATGCTGCTGATGATGATGATGATGATGATGATGATGATG 801
 Db 805 TGCCACAGGCTTACATGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 861
 QY 802 GACACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 861
 Db 862 GAGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 921
 QY 862 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 921
 Db 922 CGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 981
 QY 922 ATGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 981
 Db 982 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1041
 QY 982 ATGAGAGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1041
 Db 1042 ATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1101
 QY 1042 CGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1101
 Db 1102 GTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1161
 QY 1102 TACATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1161
 Db 1162 TACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1221
 QY 1162 CCCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1221
 Db 1222 CCCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1281
 QY 1222 GCGAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1269
 Db 1282 GCCAGCTCTGTACTGAGGCTCACCGCTCTTTGTAGGGGCTTACACCTTC 1329
 RESULT 15
 ABO30655 1536 bp mRNA linear PRI 25-MAR-2000
 LOCUS ABO30655
 DEFINITION Homo sapiens mRNA for mutant p53 binding protein 1 (MBP1), complete cds.
 ACCESSION ABO30655
 VERSION ABO30655.1 GI:7328920
 KEYWORDS MBP1; mutant p53 binding protein 1 (MBP1).
 SOURCE Homo sapiens cDNA to mRNA.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Tanaka, S., Sugimachi, K. and Sugimachi, K.
 TITLE Human mutant p53 binding protein (MBP1)
 JOURNAL Unpublished
 2 (bases 1 to 1536)
 REFERENCE Tanaka, S.
 AUTHORS Direct Submission
 TITLE Submitted (30-JUL-1999) Shinji Tanaka, Kyushu University, Faculty of
 Medicine, Department of Surgery II, 3-1-1 Maidashi, Higashi-ku,
 Fukuoka 812-8582, Japan (E-mail:shinjit@sur2.med.kyushu-u.ac.jp).

Query Match	Best Local Similarity	82.3%;	Score 1045;	DB 10;	Length 2304;
Matches 1129;	Conservative	89.0%;	Pred. No. 4.7e-289;	Indels 0;	Gaps 0;
QY	1	CAGTCACGAAATGGCTTGACCTGGATGGCCAGTCAGACACTGTTTGATATTGATGAA	60		
Db	228	CAGTCACCAACAGGCTTTGACCTGGACCGACCGCCACGACAGACAGCTGTTTGAAATATGATGAA	287		
QY	61	TGCCAACCAATCCCCGAGGCGCTGCCGAGAGACATGATGTGTGTTAAACCAATATGGCGGG	120		
Db	288	TGTGGAGCAATCCCTGAGGCTTGGCGTGGGACATGATGTGTGTCACCAACCAATATGGCGGG	347		
QY	121	TATTTATCATCTCCCGGACAAACCTGTGTATGAGAGGCGCCCTACTGAAACCCCTACTCG	180		
Db	348	TATCTGTGCATCCCTTCGAACCAACCCAGCTGTATGAGAGGCGCCCTACTGAAATCCCTACTCT	407		
QY	181	ACCCCTCACTGAGGTCCGTACCAAGCAAGCTCCCCACACACTCTAGAGTCGCAACTATATCCC	240		
Db	408	ACATCTCACTACGAGGCCCATACCGACGAGCCGACACACAGTGCAGGCTTCCAACTACCCC	467		
QY	241	ACGATCTCCAGGCGCTCTTATATATGCCGCTTTGGATATACAGATGATGATGAAGCAACCAATGT	300		
Db	468	ACGATTTTCCAGGCGCTCTTGTCTGTGCGCTTTGGGATATCAGATGATGATGAAGCAACCACTGT	527		

QY	301	GTGATGTGGACGAGTGTGGCAACAGATTCCACAGTGGACACCCACCCAGATCTGCATC	360
Db	528	TTGGATGTGGACGAGTGTGGCAACAGATTCCACAGTGGACACCCACCCAGATCTGCATC	111
QY	361	AATACGTAGAGCGGGTACACCTGCTCCGACACCGAGATATGGCTTCTGGAGGCCAG	420
Db	588	AACACGGAGAGGGGTACACCTGCTCCGACACTGATGGTACTGGCTTCTGGAGGGCGAG	647
QY	421	TGCTTAACATGTGATGAATGTGCGTATGGTTACTGTCACAGACCTGTGCGAATGTGCT	480
Db	648	TGCTTAACATGTGATGAATGTGCGTATGGTTACTGTCACAGACCTGTGCGAATGTGCT	707
QY	481	GGATCCATTTCTTTACATGCACCCGTTTACCCGATAGAGATGGAGGTCCTTC	540
Db	708	GGATCCATTTCTTTACATGCATACCCGTTTACCCGATAGAGATGGAGGTCCTTC	767
QY	541	CAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCGCAAACTGCGTCAACACTAC	600
Db	768	CAAGATGTGAACGAGTGTGAACCTGGAACCCCTGCTTTCAGACCTGCGTCAACACTAC	827
QY	601	GGCTCTTTCATCTGCCGCTGTGACCCAGATATGAACTGTGGAAATGGCGTTATTC	660
Db	828	GGTCTTTTCATCTGCCGCTGTGACCCAGATATGAACTGTGGAAATGGCGTTATTC	887
QY	661	AGTGAATATGACAGAGTGCACACTTCTCTGAGTTCTCTGCAACATGAGTGTGTAACCG	720
Db	888	AGTGAATATGAGTATGAGTGCACACTTCTCGAATCTCTCTGTCAACATGAGTGTGTAACCG	947
QY	721	CCCGGCATACTTCTGCTCCTCCGTCACAGGCTACATCTCGCGGATGAGAACCGAGC	780
Db	948	CCGGGCTCATCTTCTGCTCATGCCCTCCAGGGTACGTCCTTGTGGAAAGTAAACCGAGC	1007
QY	781	TGCCAAGATCAAGATGTGAGACAGAACACACAGTGCACACTGTCAGACAGCAGTC	840
Db	1008	TGCCAGGATATCAATGAATGTGAGACCGGAACACACAGTGCACACTGCCCTGTCAGACAGTC	1067
QY	841	TACAATTTTCAAGGGGGCTTCAAAATGCATGACCCCATCGCTGTGAGAGGCTTATCTG	900
Db	1068	TACAATTTTCAAGGGGGCTTCAAAATGTATCGAACCCCATCGCTGTGAGAGGCTTATCTG	1127
QY	901	AGGATCAGTATGAACCGCTGTATGTTGTCCTGCGAGAACCCCTGGCTGTCAGACACGCC	960
Db	1128	CTGATTTGGGAGTAAACCGCTGTATGTTGTCCTGCGAGAACCTGGCTGTCAGACACGCC	1187
QY	961	TTTACATCTTGTACCGGGACATGAGACGTGTGTGAGAGCGTCCGTCCTCCGCTGCATC	1020
Db	1188	TTTACATCTTGTGTTCGGGACATGAGATGTGTATCGAGAGCGTCTGTCCTGCTGCATC	1247
QY	1021	TTTCAAAATGCAAGCCACGACCCGCTTACCTGTGGGGCTTATTAATATTTTCCAGATCAATCT	1080
Db	1248	TTTCAAAATGCAAGCCACGACCCGCTTACCTGTGGGGCTTATTAATATTTTCCAGATCAATCT	1307
QY	1081	GGGATATGAGGCGAGAGATTTTACATGCGGCAAAAGGGGCCCATCATGATGCCACCTGTGTG	1140
Db	1308	GGGATATGAGGCTGAGAGATTTTACATGCGGCAAAAGGGGCCCATCATGATGCCACCTGTGTG	1367
QY	1141	ATGACAGCCCCATCAAAAGGCGCCCGGGAAATTCACGTGTGGACTTGGAAATGATCACTGTC	1200
Db	1368	ATGACAGCCCCATCAAAAGGCGCTTCGGGACATCCAGCTGGAAATGGAAGTATCAACCGTC	1427
QY	1201	AACACTGTCAATCAACTTCAGAGCAGCTCGGTGATCCAGATGCGGATATATTTGTGCGAG	1260
Db	1428	AACACTGTCAATCAACTTCAGAGCAGCTCGGTGATCCAGATGCGGATATATTTGTGCGAG	1487
QY	1261	TACCAATTC 1269	
Db	1488	TATCGGTTTC 1496	

RESULT 14	
HSAL32819	
LOCUS	HSAL32819
DEFINITION	Homo sapiens mRNA for fibulin-4.
	1440 bp
	mRNA
	linear
	PRI 15-DEC-2000

RESULT 12
 AF112153
 LOCUS
 DEFINITION Rattus norvegicus developmental arteries and neural crest EGF-like
 protein mRNA, complete cds.
 ACCESSION AF112153
 VERSION AF112153.1 GI:5305674
 KEYWORDS
 SOURCE
 ORGANISM Rattus norvegicus.
 Rattus norvegicus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 2499)
 AUTHORS Nakamura, T., Ruiz-Lozano, P., Lindner, V., Yabe, D., Tanikawa, M.,
 Furukawa, Y., Kobuke, K., Tashiro, K., Lu, Z., Andon, N.L., Schaub, R.,
 Matsumoto, A., Sasayama, S., Chien, K.R. and Honjo, T.
 DANCE, a novel secreted RGD protein expressed in developing,
 atherosclerotic, and balloon-injured arteries
 J. Biol. Chem. 274 (32), 22476-22483 (1999)
 JOURNAL 99357779
 MEDLINE 10428822
 PUBMED 2 (bases 1 to 2499)
 REFERENCE Nakamura, T., Yabe, D., Tashiro, K. and Honjo, T.
 AUTHORS Nakamura, T., Yabe, D., Tashiro, K. and Honjo, T.
 TITLE Direct Submission
 JOURNAL Submitted (09-DEC-1998) Medicine, University of California San
 Diego, 9500 Gilman Drive, La Jolla, CA 92093-0613, USA
 FEATURES
 source Location/Qualifiers
 1..2499
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 350..1696
 /note="DANCE: contains 6 cbEGF domains and 1 RGD motif;
 secreted protein; expressed in developmental aorta, neural
 crest cells, balloon injured vessels, atherosclerotic
 lesions"
 /codon_start=1
 /product="developmental arteries and neural crest BGF-like
 protein"
 /protein_id="AA041769.1"
 /db_xref="GI:5305675"
 /translation="MPEGLKRLIVYTIALLMPLHPNAAQQCTNGFDLDROTCIDID
 ERTIPEACRDMACVONGYLCIPRTNPNVRYGPNYSYSGPPAAPYPAS
 NPPTISRPVLCRGYOMDEGNOCVDDECAHSDHCNPOTICIMTEGGYCTGNGY
 LLEGCLDIDECRYGCOOLCANVPGSYSCNRPGLINDGRSQDNEPTEPCV
 QTQVNTYGFICRDPGYELEEDGSHCSDMDECSFSEFLCOHECVNGSTFSCPPG
 YVLEENRSCQDINECEHNRHNTCTPILOTVCYNLOGFKCIDIVEEPYLLIGDNRCMC
 PAENITGCRDQPTILFRMDVDSGRVADIFQMAATRYGAYVIFQIKSGNEGREF
 VMKRGPSIATLVMTRPKGRPDIDLEMITVNTVINFGRSSVYRLRIYVSQYF"
 BASE COUNT 584 a 669 c 595 g 649 t 2 others
 ORIGIN
 Query Match 82.5%; Score 1046.6; DB 10; Length 2499;
 Best Local Similarity 89.0%; Pred. No. 1.6e-289;
 Matches 1130; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

241 ACATCTCCAGGCTCTTATATGCCGCTTGGATACAGATGATGAAGCAACCAATGT 300
 665 ACATTTCCAGGCTCTTGTCTGTGCGTTTGGATATGATGATGAAGCAACCAATGT 724
 301 GTGGATGTGACAGATGTGCACAGATTTCCACAGTCAACCCACCAAGATGTGCATC 360
 725 GTGGATGTGACAGATGTGCACAGATTTCCACAGTCAACCCACCAAGATGTGCATC 784
 361 AATACGAAGGGGTACCTGCTGCTGACCCACAGGATATGTGCTTGAAGGCGAG 420
 785 AACACGGAAGGAGGTACCTGCTGCTGACCTGATGAGGTGCTTGAAGGCGAG 844
 421 TGCCTAGATATGTATGATATGCTGATGCTGATGCTGACAGCTGTGCAATGTTCT 480
 845 TGCCTAGATATGTATGATATGCTGATGCTGATGCTGACAGCTGTGCAATGTTCT 904
 481 GGATCTATTTCTGATACATGCAACCTGCTGTTTACCTTCATGATGATGAAGTCTGC 540
 905 GGATCTATTTCTGATACATGCTGATGCTGCTGCTTCAACCATGATGGAAGTCTGC 964
 541 CAAGATGTGAACGAGTGTGCGACGAGAACCCCTGCGTCAACCTGGGTCAACCTTAC 600
 965 CAAGATGTGAACGAGTGTGGAACCTGAGAACCCCTGCTGCTGCTGCTGCTGCTGCT 1024
 601 GGCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 1025 GGCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1084
 661 AGTATATGACAGATGACAGTCTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 1085 AGTATATGACAGATGACAGTCTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1144
 721 CCGGCACTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 1145 CCGGCACTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1204
 781 TGCCAGACATCAACGATGTGACACAGAACCAACAGACAGTCAACCTGACAGACAGTGC 840
 1205 TGCCAGATATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1264
 841 TACATTTACAGGGGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 900
 1265 TACATTTACAGGGGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1324
 901 AGATATGATATCAACGATGTGACACAGAACCAACAGACAGTCAACCTGACAGACAGTGC 960
 1325 CTGATTTGGATATCAACGATGTGACACAGAACCAACAGACAGTCAACCTGACAGACAGTGC 1384
 961 TTTACATCTTGTACCGGACATGAGAGTGTGATGAGAGAGTGTGATGAGAGAGTGTGAT 1020
 1385 TTTACATCTTGTACCGGACATGAGAGTGTGATGAGAGAGTGTGATGAGAGAGTGTGAT 1444
 1021 TTCCAAATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
 1445 TTCCAAATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1504
 1081 GGGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
 1505 GGGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1564
 1141 ATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
 1565 ATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1624
 1201 AACACTGTATCAACTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
 1625 AACACTGTATCAACTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1684
 1261 TACCATTC 1269
 1685 TATCCGTT 1693


```

RESULT 9
AF093118
LOCUS Homo sapiens UP50 mRNA, complete cds.
ACCESSION AF093118
VERSION AF093118.1 GI:3676821
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2019)
AUTHORS Zemel, R., Sholtz, O. and Shaul, Y.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-1998) Dept. of Molecular Genetics, Weizmann
Institute of Science, Rehovot 76100, Israel
FEATURES
Source
1..2019
/organism="Homo sapiens"
/db_xref="taxon:9606"
/issue_type="urine"
5'UTR
1..168
169..1515
/note="50kD; similar to human SI-5 encoded by GenBank
Accession Number U03877; possible membrane and
extracellular matrix protein; urine p50 protein"
CDS
169..1515
/codon_start=1
/product="p50"
/protein_id="AAC2107.1"
/db_xref="GI:3676822"
/translation="MPGIRKILVTIILALCLPSPNAOQCTNGFDLDRSGOCLDID
ECHTIPKACGDMVQNGYLCVDSRTNIPYRGPVSNIPYSPYGPAPAPLISAP
NYPTISPLICRFYQDMDSNOCVDEQVDSHOCNPOTICIMKGGYSCSDGYV
LLEGOCLEIDIEGRGYCOOLCANYPGVSCTCNBGRFNLNDRSCODVNECATENPCV
QGVNTYGSFICRCDPEYELRDEGVHVSQSDMECSFSEFLCOHECVQGVYFCSPG
YILLDRSCODINECHRNHTCRLQDTGICNLQGGFCIDIRCEPEYLRISDNKCMC
PAENPCRDQPFILIRDMDVSVSRVPAIDFQMAATRYPGAVYIIOIKSGNREF
YMRQGTISATVLTIRPIKGRREIQLDLEMITVNTVINFRSVIRLRITYSQPF"
1516..2019
BASE COUNT 478 a 585 c 490 g 465 t 1 others
ORIGIN
3'UTR
Query Match 98.7%; Score 1253; DB 9; Length 2019;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1259; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Y 1 CAGTGCAGCAATGGCTTTGACCTGATCGCAGTCAGACAGTGTAGATATGATGAA 60
|||||
Db 244 CAGTGCAGCAATGGCTTTGACCTGATCGCAGTCAGACAGTGTAGATATGATGAA 303
61 TCCCGAACCATCCCGAGGCGCTGCCGAGAGACATGATGTGTTAACCAAAATGGCGG 120
|||||
Db 304 TCCCGAACCATCCCGAGGCGCTGCCGAGAGACATGATGTGTTAACCAAAATGGCGG 363
121 TATTATGATTCGCCGGAACAAACCTGTGTATCGAGGGCGCTACTGAAACCCCTACTCG 180
|||||
Db 364 TATTATGATTCGCCGGAACAAACCTGTGTATCGAGGGCGCTACTGAAACCCCTACTCG 423
181 ACCCCCTACTCAGTCCGTACCCAGAGCTGCCCACTCAGTCCAGTCCAACTATACCC 240
|||||
Db 424 ACCCCCTACTCAGTCCGTACCCAGAGCTGCCCACTCAGTCCAGTCCAACTATACCC 483
241 ACGATCTCCAGGCGCTTATATGCGCTTTGGATACCAAGATGATGAAAGCAACAATGT 300
|||||
Db 484 ACGATCTCCAGGCGCTTATATGCGCTTTGGATACCAAGATGATGAAAGCAACAATGT 543
301 GTGATGTGAGAGATGTGCAACAGATTCGCCACAGTCGAACCCCAACAGATGTGCATC 360
|||||
Db 544 GTGATGTGAGAGATGTGCAACAGATTCGCCACAGTCGAACCCCAACAGATGTGCATC 603
361 AATACGAGGCGGTACACCTGCTCTGACCGAGCATATTTGGTTGGAAAGGCGAG 420
|||||

```

```

Db 604 AATATGAGGCGGTACACCTGCTCCGACCCGAGATATGGCTTTGGAAAGCCAG 663
QY 421 TGGTTAGACATTTGATGATGTCGCTATGTTACTGCGCAGCAGCTGTGCCAATCTCT 480
|||||
Db 664 TGGTTAGACATTTGATGATGTCGCTATGTTACTGCGCAGCAGCTGTGCCAATCTCT 723
QY 481 GGATCTTATCTTGTATACATGCAACCTGGTTTATCCCTCAATGAGATGAAAGTCTTGC 540
|||||
Db 724 GGATCTTATCTTGTATACATGCAACCTGGTTTATCCCTCAATGAGATGAAAGTCTTGC 783
541 CAAGATGTGAACGATGTGCAACCGAGAACCCCTGGCTGCAAACTGTGCTACACCTTAC 600
|||||
Db 784 CAAGATGTGAACGATGTGCAACCGAGAACCCCTGGCTGCAAACTGTGCTACACCTTAC 843
QY 601 GGCTCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGAAATGGCTTCATTTGC 660
|||||
Db 844 GGCTCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGAAATGGCTTCATTTGC 903
QY 661 AGTATATGAGACAGTGCAGCTTCTGTGAGTTCCTTGCCAAACATGAGTGTGAAACCAG 720
|||||
Db 904 AGTATATGAGACAGTGCAGCTTCTGTGAGTTCCTTGCCAAACATGAGTGTGAAACCAG 963
721 CCCGGACATACTTCTGCTCCTGCTCCAGGCTACATCTCTGCTGATATGACACCGAAGC 780
|||||
Db 964 CCCGGACATACTTCTGCTCCTGCTCCAGGCTACATCTCTGCTGATATGACACCGAAGC 1023
QY 781 TGCAGAGACATCAACGAATGTGAGACAGAGAACACAGTGCACCAACCTGACAGACCTGC 840
|||||
Db 1024 TGCAGAGACATCAACGAATGTGAGACAGAGAACACAGTGCACCAACCTGACAGACCTGC 1083
841 TGCATTTTCAAGGGGGCTTTCAATGATCAGACCCCATCCGCTGTGAGAGACCTTATCTG 900
|||||
Db 1084 TGCATTTTCAAGGGGGCTTTCAATGATCAGACCCCATCCGCTGTGAGAGACCTTATCTG 1143
QY 901 AGATATGATTAACCGCTGTATGTCTGTCTGTGAGAACCCCTGGCTGACAGACAGCCG 960
|||||
Db 1144 AGATATGATTAACCGCTGTATGTCTGTCTGTGAGAACCCCTGGCTGACAGACAGCCG 1203
961 TTTACCATTTTACCGGAGCATGAGCGTGTGTCAAGAGCTCCGTTCCGCTGACATC 1020
|||||
Db 1204 TTTACCATTTTACCGGAGCATGAGCGTGTGTCAAGAGCTCCGTTCCGCTGACATC 1263
QY 1021 TTCCAAATGCAAGCCACGACCGCTACCTCGGGGGCTATTTACATTTTCCAGATCAATCT 1080
|||||
Db 1264 TTCCAAATGCAAGCCACGACCGCTACCTCGGGGGCTATTTACATTTTCCAGATCAATCT 1323
QY 1081 GGGATATGAGGCGAGAAATTTTACATGCGGCAAAACGGGCCCATCATGTCACCCCTG 1140
|||||
Db 1324 GGGATATGAGGCGAGAAATTTTACATGCGGCAAAACGGGCCCATCATGTCACCCCTG 1383
1141 ATGACAGCGCCCATCAAGAGGCGCCCGGAAATCCAGCTGGACTTGGAAATGATCACTGTC 1200
|||||
Db 1384 ATGACAGCGCCCATCAAGAGGCGCCCGGAAATCCAGCTGGACTTGGAAATGATCACTGTC 1443
QY 1201 AACACTGTCACTCAACTTCAAGAGGAGCTCCGATCCGACTCGGATATATGTGTCGAG 1260
|||||
Db 1444 AACACTGTCACTCAACTTCAAGAGGAGCTCCGATCCGACTCGGATATATGTGTCGAG 1503
QY 1261 TACCATTTTC 1269
|||||
Db 1504 TACCATTTTC 1512

```

```

RESULT 10
BC006636
LOCUS BC006636
DEFINITION Mus musculus, fibulin 5, clone MGC:5656 IMAGE:3482574, mRNA,
complete cds.
ACCESSION BC006636
VERSION BC006636.1 GI:13879321
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus

```

QY 961 TTATACATCTTGTACCGGAGATGAGAGCTGTGTACAGAGCTCCGTTCCCGTGAATC 1020
 DB 1485 TTATACATCTTGTACCGGAGATGAGAGCTGTGTACAGAGCTCCGTTCCCGTGAATC 1544
 QY 1021 TTCCAAATGCAAGCCAGACCGCTTACCTGGGGCTTATTAATTTTCAGATCAAAATCT 1080
 DB 1545 TTCCAAATGCAAGCCAGACCGCTTACCTGGGGCTTATTAATTTTCAGATCAAAATCT 1604
 QY 1081 GGAATGAGGGGAGAAATTTTACATGCGGGGAAAGGGGGCCCATCAGAGCCACCTGTG 1140
 DB 1605 GGAATGAGGGGAGAAATTTTACATGCGGGGAAAGGGGGCCCATCAGAGCCACCTGTG 1664
 QY 1141 ATGACAGCGCCCATCAAGAGGCGCCGGGAAATCCAGCTGAGCTTGAATATCACTGTC 1200
 DB 1665 ATGACAGCGCCCATCAAGAGGCGCCGGGAAATCCAGCTGAGCTTGAATATCACTGTC 1724
 QY 1201 AACACTGTATCACTTCAGAGGAGCTCCGTGATCCGACTGGGATATATGTGTGAG 1260
 DB 1725 AACACTGTATCACTTCAGAGGAGCTCCGTGATCCGACTGGGATATATGTGTGAG 1784
 QY 1261 TACCCATTC 1269
 DB 1785 TACCCATTC 1793

RESULT 8
 AX403659 2609 bp DNA linear PAT 14-JUN-2002
 LOCUS Sequence 14 from Patent W00077037.
 DEFINITION AX403659
 ACCESSION AX403659 GI:21437116
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
 AUTHORS
 1
 Ashkenazi, A., Baker, K., Botstein, D., Desnoyers, L., Eaton, D.L.,
 Ferrara, N., Fong, S., Gao, W.Q., Gerber, H., Gerritsen, M.E.,
 Goddard, A., Godowski, P., Gurney, A., Kijavich, I.J., Mather, J.,
 Napier, M., Pan, J., Paoni, N., Roy, M., Tumas, D., Watanabe, C.,
 Williams, P.M., Wood, W.I. and Zhang, Z.
 Secreted and transmembrane polypeptides and nucleic acids encoding
 the same
 JOURNAL Patent: WO 0077037-A 14 21-DEC-2000;
 Genentech Inc. (US)

FEATURES
 Source
 Location/Qualifiers
 1..2609
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 622 a 718 c 622 g 647 t
 ORIGIN

Query Match 99.9%; Score 1267.4; DB 6; Length 2609;
 Best Local Similarity 99.9%; Pired. No. 0;
 Matches 1268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAGTGCAGCAATGGCTTTGACCTGATGCGCAGTCAAGAGAGCTTTATGATATGATGAA 60
 DB 525 CAGTGCAGCAATGGCTTTGACCTGATGCGCAGTCAAGAGAGCTTTATGATATGATGAA 584
 QY 61 TCCCGAAGCATCCCGAGGCTCCCGAGAGACATGATGTGTAAACCAAAATGGCGG 120
 DB 585 TCCCGAAGCATCCCGAGGCTCCCGAGAGACATGATGTGTAAACCAAAATGGCGG 644
 QY 121 TATTTATGATTCCTCCGAGCAAAACCTGTATGAGAGGGCCCTACTGAAACCCCTACTG 180
 DB 645 TATTTATGATTCCTCCGAGCAAAACCTGTATGAGAGGGCCCTACTGAAACCCCTACTG 704
 QY 181 ACCCCCTACTCAGGTCCGTACCAAGAGAGCTGCCCCACACACTCTAGCTCAAAATATCCC 240
 DB 705 ACCCCCTACTCAGGTCCGTACCAAGAGAGCTGCCCCACACACTCTAGCTCAAAATATCCC 764

QY 241 ACGATCTCCAGGCTCTTATATGCGGCTTGGATACAGATGATGAAGAACAACATGT 300
 DB 765 ACGATCTCCAGGCTCTTATATGCGGCTTGGATACAGATGATGAAGAACAACAATGT 824
 QY 301 GTGGATGTGAGCAGATGTGCAACAGATTCACCAGTGCACACCCACAGATCTGCATC 360
 DB 825 GTGGATGTGAGCAGATGTGCAACAGATTCACCAGTGCACACCCACAGATCTGCATC 884
 QY 361 AATACGGAAGCGGGTACACTGCTCTGCAACGAGATATTTGGCTTCTGGAAGCCAG 420
 DB 885 AATACGGAAGCGGGTACACTGCTCTGCAACGAGATATTTGGCTTCTGGAAGCCAG 944
 QY 421 TCGTTAGACATTTGATGATTCGCTATGATGTTACTGCGCAGAGCTGTGCGAATTTCT 480
 DB 945 TCGTTAGACATTTGATGATTCGCTATGATGTTACTGCGCAGAGCTGTGCGAATTTCT 1004
 QY 481 GGATCTATTTCTGTACATCAACCCCTGGTTTACCTCAATGAGAGTGAAGGTCTTGC 540
 DB 1005 GGATCTATTTCTGTACATCAACCCCTGGTTTACCTCAATGAGAGTGAAGGTCTTGC 1064
 QY 541 CAAGATGTGACAGAGTGTGCAACGAGAACCCCTGCGCAACCTGGCTCAACACTAC 600
 DB 1065 CAAGATGTGACAGAGTGTGCAACGAGAACCCCTGCGCAACCTGGCTCAACACTAC 1124
 QY 601 GGCTCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGAGATGAGGCTTCAATGC 660
 DB 1125 GGCTCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGAGATGAGGCTTCAATGC 1184
 QY 661 AGTATATGAGCAGAGTGTGCTGTGAGTCTCTGCAACATGAGTGTGAACAG 720
 DB 1185 AGTATATGAGCAGAGTGTGCTGTGAGTCTCTGCAACATGAGTGTGAACAG 1244
 QY 721 CCGGAGACATCTTCTGCTCTGCTGCGTCAAGAGTACATCTGCTGAGATGAGCAACGAGC 780
 DB 1245 CCGGAGACATCTTCTGCTCTGCTGCGTCAAGAGTACATCTGCTGAGATGAGCAACGAGC 1304
 QY 781 TGCAGAGACATCAACGATGTGAGCAGAGAACACAGCTGCAACCTGACAGAGAGCTGC 840
 DB 1305 TGCAGAGACATCAACGATGTGAGCAGAGAACACAGCTGCAACCTGACAGAGAGCTGC 1364
 QY 841 TACAAATTTACAAAGGGGCTTCAAAATGATGATGACCCCATCCGCTGTGAGAGGCTTATCTG 900
 DB 1365 TACAAATTTACAAAGGGGCTTCAAAATGATGATGACCCCATCCGCTGTGAGAGGCTTATCTG 1424
 QY 901 AGGATCAGTATACCGCTGTATGTCTGCTGAGAACCTTGCTGAGAGAGCAGGCC 960
 DB 1425 AGGATCAGTATACCGCTGTATGTCTGCTGAGAACCTTGCTGAGAGAGCAGGCC 1484
 QY 961 TTATACATCTTGTACCGGAGATGAGAGTGTGTACAGAGCTCCGTTCCCGTGAATC 1020
 DB 1485 TTATACATCTTGTACCGGAGATGAGAGTGTGTGTACAGAGCTCCGTTCCCGTGAATC 1544
 QY 1021 TTCCAAATGCAAGCCAGACCGCTTACCTGGGGCTTATTAATTTTCAGATCAAAATCT 1080
 DB 1545 TTCCAAATGCAAGCCAGACCGCTTACCTGGGGCTTATTAATTTTCAGATCAAAATCT 1604
 QY 1081 GGAATGAGGGGAGAGAAATTTTACATGCGGGGAAAGGGGGCCCATCAGAGCCACCTGTG 1140
 DB 1605 GGAATGAGGGGAGAGAAATTTTACATGCGGGGAAAGGGGGCCCATCAGAGCCACCTGTG 1664
 QY 1141 ATGACAGCGCCCATCAAGAGGCGCCGGGAAATCCAGCTGAGCTTGAATATCACTGTC 1200
 DB 1665 ATGACAGCGCCCATCAAGAGGCGCCGGGAAATCCAGCTGAGCTTGAATATCACTGTC 1724
 QY 1201 AACACTGTATCAACTTCAGAGGAGCTCCGTATGCGATGCGAGATATATGTGTGCGAG 1260
 DB 1725 AACACTGTATCAACTTCAGAGGAGCTCCGTATGCGATGCGAGATATATGTGTGCGAG 1784
 QY 1261 TACCCATTC 1269
 DB 1785 TACCCATTC 1793

FEATURES	SEQUENCE	LOCATION	FUNCTION	REFERENCE
QY	541	CAAGATGGAAGAGTGTGGACCGAGAGAACCCCTGGCGGAAACCTGGCTGTAACACCTAC	600	
Db	997	CAGATGTGAAGAGATGTGCCACCGAGAACCCCTGGCGGAAACCTGGCTGTAACACCTAC	1056	
QY	601	GGCTTTTTCATCTGGCCGCTGTGACCCAGATATGAACTTGAGGAAGATGGCTTCATTGC	660	
Db	1057	GGCTTTTTCATCTGGCCGCTGTGACCCAGATATGAACTTGAGGAAGATGGCTTCATTGC	1116	
QY	661	AGTATATGAGAGAGTGCAGCTTCTCTGTAGTTCTCTGTCCAAATGATGTGTGACCG	720	
Db	1117	AGTATATGAGAGAGTGCAGCTTCTCTGTAGTTCTCTGTCCAAATGATGTGTGACCG	1176	
QY	721	CCGGGACATATCTTCTGCTCTGCTCCCTCAGAGCTACATCTGCTGTGAGTGAACAACGAGC	780	
Db	1177	CCGGGACATATCTTCTGCTCTGCTCCCTCAGAGCTACATCTGCTGTGAGTGAACAACGAGC	1236	
QY	781	TGCCAAGACATCAAGAAATGTGTAGACACAGAAACACAGCTGCMACTCTGACAGACGTGC	840	
Db	1237	TGCCAAGACATCAAGAAATGTGTAGACACAGAAACACAGCTGCMACTCTGACAGACGTGC	1296	
QY	841	TACAATTTACAAAGGGGGCTTCAAAATGCATGCAGACCCATCCGCTGTGAGAGACCTTATCTG	900	
Db	1297	TACAATTTACAAAGGGGGCTTCAAAATGCATGCAGACCCATCCGCTGTGAGAGACCTTATCTG	1356	
QY	901	AGGATCACTGATTAACCCGCTGTATGTCTCTGTGAGAAACCTGTGCTGCAGAGACAGCC	960	
Db	1357	AGGATCACTGATTAACCCGCTGTATGTCTCTGTGAGAAACCTGTGCTGCAGAGACAGCC	1416	
QY	961	TTTACCATCTTGTACCCGGGACATGAGACGTGGTGTGACAGACCTCCCTCCCGCTGACATC	1020	
Db	1417	TTTACCATCTTGTACCCGGGACATGAGACGTGGTGTGACAGACCTCCCTCCCGCTGACATC	1476	
QY	1021	TTTCCAAATGCAAGCCACACCCCGCTACCCCTGGGGCTTATACATTTTTCAGATCAAAATCT	1080	
Db	1477	TTTCCAAATGCAAGCCACACCCCGCTACCCCTGGGGCTTATACATTTTTCAGATCAAAATCT	1536	
QY	1081	GGGAATGAGGGCAGAGAAATTTTACATGTGGGCAAAACGGGCCCCATCAGTGTGCCACCTGGTG	1140	
Db	1537	GGGAATGAGGGCAGAGAAATTTTACATGTGGGCAAAACGGGCCCCATCAGTGTGCCACCTGGTG	1596	
QY	1141	ATGACACGCCCCATCAAAAGGGCCCCGGGAAATTCAGCTGTGAGCTTGGAAATCATCATGCTC	1200	
Db	1597	ATGACACGCCCCATCAAAAGGGCCCCGGGAAATTCAGCTGTGAGCTTGGAAATCATCATGCTC	1656	
QY	1201	AACACTGTTCATCAACTTCAGAGGACGTCCGTGTATCGCATGCGGATATATGTGTCCGAG	1260	
Db	1657	AACACTGTTCATCAACTTCAGAGGACGTCCGTGTATCGCATGCGGATATATGTGTCCGAG	1716	
QY	1261	TACCCATTTC 1269		
Db	1717	TACCCATTTC 1725		
RESULT 7				
LOCUS	AX376340	2609 bp	DNA	linear
DEFINITION	Sequence 407 from Patent WO0168848.			PAT 01-MAR-2002
ACCESSION	AX376340			
VERSION	AX376340.1	GI:19170561		
KEYWORDS	human.			
SOURCE	Homo sapiens			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE	1	Baker, K.P., Chen, J., Desnoyers, L., Goddard, A., Godowski, P.J., Guney, A.L., Pan, J., Smith, V., Watanabe, C.K., Wood, W.I. and Zhang, Z.		
TITLE	Secreted and transmembrane polypeptides and nucleic acids encoding the same			
JOURNAL	Patent: WO 0168848-A 407 20-SEP-2001; Genentech, Inc. (US)			
FEATURES	Location/Qualifiers			

Source	1. 2609	"/organism="Homo sapiens"	"/db_xref="taxon:9606"
BASE COUNT	622 a	718 c	622 g 647 t
ORIGIN			
Query Match	99.9%	Score 1267.4	DB 6; Length 2609;
Best Local Similarity	99.9%	Pred. No. 0;	
Matches 1268; Conservative	0;	Mismatches	1; Indels 0; Gaps 0;
OY	1	CAGTGCACGATGCGTTTGACCTGGATGCGCAGTCAGACAGTGTGTTGATATTGATTA	60
DB	525	CAGTGCACGATGCGTTTGACCTGGATGCGCAGTCAGACAGTGTGTTGATATTGATTA	584
OY	61	TGCCGAACCATCCCGAGGCGCTGCCGAGAGACATGATGTGTGTTAACCAAAATGGCGG	120
DB	585	TGCCGAACCATCCCGAGGCGCTGCCGAGAGACATGATGTGTGTTAACCAAAATGGCGG	644
OY	121	TATTTATGATATCCCGGACAAACCTGTGTATGAGGGGCCCTACTGAAACCCCTACTCG	180
DB	645	TATTTATGATATCCCGGACAAACCTGTGTATGAGGGGCCCTACTGAAACCCCTACTCG	704
OY	181	ACCCGCTACTAGGCGCTTACCCAGACAGTCCGACACCTGACCTGACCTGCAACATATCC	240
DB	705	ACCCGCTACTAGGCGCTTACCCAGACAGTCCGACACCTGACCTGACCTGCAACATATCC	764
OY	241	ACGATCTCCAGGCGCTTATATATGCGCGCTTTGGATACCAAGATGATGAAAGCAACCAATGT	300
DB	765	ACGATCTCCAGGCGCTTATATATGCGCGCTTTGGATACCAAGATGATGAAAGCAACCAATGT	824
OY	301	GTGATGTGACGAGTGTGCAACAGATTCCACACATGCAACCCACCAGATCTGCATC	360
DB	825	GTGATGTGACGAGTGTGCAACAGATTCCACACATGCAACCCACCAGATCTGCATC	884
OY	361	AATACTGAAGCGGGGTACACCTGCTCCGCGACGACGAGATTTGGCTTCTGGAAGCGCAG	420
DB	885	AATACTGAAGCGGGGTACACCTGCTCCGCGACGACGAGATTTGGCTTCTGGAAGCGCAG	944
OY	421	TGCTTAGACATTTGATGATGTGCGTATGTTACTGCGACAGCTCTGTGCAATGTTCT	480
DB	945	TGCTTAGACATTTGATGATGTGCGTATGTTACTGCGACAGCTCTGTGCAATGTTCT	1004
OY	481	GGATCTTATTTGTGACATGCAACCCGTGTTTACCTCAATGAGAGATGGAAGTCTTGC	540
DB	1005	GGATCTTATTTGTGACATGCAACCCGTGTTTACCTCAATGAGAGATGGAAGTCTTGC	1064
OY	541	CAAGATGGAAGAGTGTGCCACGAGAAACCCGCGCGCAAACTGGTGCATACACCTAC	600
DB	1065	CAAGATGGAAGAGTGTGCCACGAGAAACCCGCGCGCAAACTGGTGCATACACCTAC	1124
OY	601	GGCTCTTTCATCTCCGCTGTGACCCAGATATTAACCTTGAGGAAGATGGCGTTTCATTTGC	660
DB	1125	GGCTCTTTCATCTCCGCTGTGACCCAGATATTAACCTTGAGGAAGATGGCGTTTCATTTGC	1184
OY	661	AGTGAATGACGAGTGCAGCTTCTCTGAGTTCTCTGCCAATGAGTGTGTAACACG	720
DB	1185	AGTGAATGACGAGTGCAGCTTCTCTGAGTTCTCTGCCAATGAGTGTGTAACACG	1244
OY	721	CCGGGCAATCTTCTGCTCTGCGCCACGAGCAATCCGCGGAGTGAACCAACCGAAGC	780
DB	1245	CCGGGCAATCTTCTGCTCTGCGCCACGAGCAATCCGCGGAGTGAACCAACCGAAGC	1304
OY	781	TGCCAAGACATCAACGAATGTGACACAGAAACACAGTGCACATCTGCACAGACGTGC	840
DB	1305	TGCCAAGACATCAACGAATGTGACACAGAAACACAGTGCACATCTGCACAGACGTGC	1364
OY	841	TACAATTTTACAAAGGGGCTTCAAAATGATGACACCCCATCCGCTGTGAGAGACCTTATCTG	900
DB	1365	TACAATTTTACAAAGGGGCTTCAAAATGATGACACCCCATCCGCTGTGAGAGACCTTATCTG	1424
OY	901	AGGATCAGGTAAACCGCTGTATGTGTCTGTCTGAGAAACCTGGCTGCAGAGACCAAGCC	960
DB	1425	AGGATCAGGTAAACCGCTGTATGTGTCTGTCTGAGAAACCTGGCTGCAGAGACCAAGCC	1484

```

QY 61 TGGCAACATCCCGAGGCTGCGAGGAGACATGATGTGTTAACCAATGGCGG 120
DB 517 TGGCAACATCCCGAGGCTGCGAGGAGACATGATGTGTTAACCAATGGCGG 576
QY 121 TATTATGCAATCCCGAGGAGAACCTGTGTATGAGGGCCCTACTCGAACCCCTACTG 180
DB 577 TATTATGCAATCCCGAGGAGAACCTGTGTATGAGGGCCCTACTCGAACCCCTACTG 636
QY 181 ACCCCCTACTCAGGTCCGTACCCAGCAGCTGCCACCACCTCTCAGTCCAAATATCC 240
DB 637 ACCCCCTACTCAGGTCCGTACCCAGCAGCTGCCACCACCTCTCAGTCCAAATATCC 696
QY 241 ACATGCTCCAGGCTCTTATATGCGGCTTGTGATACGATGATGAAAGCAACCATGT 300
DB 697 ACATGCTCCAGGCTCTTATATGCGGCTTGTGATACGATGATGAAAGCAACCATGT 756
QY 301 GTGGATGTGAGCAGAGTGTGCAACAGATTTCCACAGTGTCAACCCACCCAGATTCATC 360
DB 757 GTGGATGTGAGCAGAGTGTGCAACAGATTTCCACAGTGTCAACCCACCCAGATTCATC 816
QY 361 AATACGTGAAGGGGCTACCTGCTCTGCAACGAGATTTGGCTTCTGAGAGCCAG 420
DB 817 AATACGTGAAGGGGCTACCTGCTCTGCAACGAGATTTGGCTTCTGAGAGCCAG 876
QY 421 TGTGTAGACATGTGATGATGTGCTATGTGTTACTGCCAGAGCTGTGCGAATGTCTCT 480
DB 877 TGTGTAGACATGTGATGATGTGCTATGTGTTACTGCCAGAGCTGTGCGAATGTCTCT 936
QY 481 GGATCTTATTTCTGTATCATCAACCCCTGTGTTTACCTCAATGAGATGGAAGTCTTGC 540
DB 937 GGATCTTATTTCTGTATCATCAACCCCTGTGTTTACCTCAATGAGATGGAAGTCTTGC 996
QY 541 CAAGATGTGAACAGTGTGCGACCGAGAACCCCTGCGTGCAGAACTGGGTCAACACTTAC 600
DB 997 CAAGATGTGAACAGTGTGCGACCGAGAACCCCTGCGTGCAGAACTGGGTCAACACTTAC 1056
QY 601 GGCTCTTTCATCTGCCCTGTGACCCAGATATGAATGAGAGAGGGCTTATATTC 660
DB 1057 GGCTCTTTCATCTGCCCTGTGACCCAGATATGAATGAGAGAGGGCTTATATTC 1116
QY 661 AGTGATATGAGCAGAGTGTGCTGTGAGTCTCTGCCAATCATGAGTGTGTAACAG 720
DB 1117 AGTGATATGAGCAGAGTGTGCTGTGAGTCTCTGCCAATCATGAGTGTGTAACAG 1176
QY 721 CCGGACATACATCTGCTCTGCTGCTGCAAGTACATCTGCTGAGATGAGAACCGAAGC 780
DB 1177 CCGGACATACATCTGCTCTGCTGCTGCAAGTACATCTGCTGAGATGAGAACCGAAGC 1236
QY 781 TGGCAAGACATCAAGATGTGACAGAGAGAACCAACAGTGCACCTCAGCAGAGAGTGC 840
DB 1237 TGGCAAGACATCAAGATGTGACAGAGAGAACCAACAGTGCACCTCAGCAGAGAGTGC 1296
QY 841 TACATATTTACAGAGGGGCTTCAATTCATGCAACCCCATCCGCTGTGAGAGGCTTATCTG 900
DB 1297 TACATATTTACAGAGGGGCTTCAATTCATGCAACCCCATCCGCTGTGAGAGGCTTATCTG 1356
QY 901 AGGATGATGATTAACCGCTGATGTGCTGCTGAGAACCTGGGTGAGAGAGACAGGCC 960
DB 1357 AGGATGATGATTAACCGCTGATGTGCTGCTGAGAACCTGGGTGAGAGAGACAGGCC 1416
QY 961 TTTTACATCTTTTACCGGAGCATGAGAGTGTGTACAGAGAGCTCCGTTCCGCTGACATC 1020
DB 1417 TTTTACATCTTTTACCGGAGCATGAGAGTGTGTACAGAGAGCTCCGTTCCGCTGACATC 1476
QY 1021 TTTTCAATATGCAAGCAGACCCGCTACCCCTGAGGCTTATACATTTTTCAGATCAATCT 1080
DB 1477 TTTTCAATATGCAAGCAGACCCGCTACCCCTGAGGCTTATACATTTTTCAGATCAATCT 1536
QY 1081 GGGATATGAGGGCAGAGATTTTACATGCGGCAACGGGCCCCCATGAGTGCACCTGGTG 1140
DB 1537 GGGATATGAGGGCAGAGATTTTACATGCGGCAACGGGCCCCCATGAGTGCACCTGGTG 1596

```

```

QY 1141 ATGACAGCCCATCAAGAGGGGCCCCGGGAAATCCAGCTGGAGCTTGGAAATGATCACTGC 1200
DB 1597 ATGACAGCCCATCAAGAGGGGCCCCGGGAAATCCAGCTGGAGCTTGGAAATGATCACTGC 1656
QY 1201 AACACTGTCATCAACTTCAGAGGAGAGTCCGTGATCCGATCCGATATATGTGTGCGAG 1260
DB 1657 AACACTGTCATCAACTTCAGAGGAGAGTCCGTGATCCGATCCGATATATGTGTGCGAG 1716
QY 1261 TACCATTCC 1269
DB 1717 TACCATTCC 1725

RESULT 6
LOCUS AR173204 2550 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 2 from patent US 6303765.
ACCESSION AR173204
VERSION AR173204.1 GI:17912695
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 2550)
Bandman, O., Corley, N.C. and Guegler, K.J.
Human extracellular matrix proteins
Patent: US 6303765-A 2.16-OCR-2001;
Location/Qualifiers
1..2550 /organism="unknown"
BASE COUNT 623 a 690 c 596 g 641 t
ORIGIN

Query Match 100.0%; Score 1269; DB 6; Length 2550;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

FEATURES passed the following selection criteria: matched mRNA gi: 5453649.
 source Location/Qualifiers

1. 2367
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:22412 IMAGE:4693953"
 /tissue_type="Lung"
 /clone_lib="NHL MGC_77"
 /lab_host="DH10B"
 /note="Vector: pDNR-LIB"
 184. .1530
 /product="Unknown (protein for MGC:22412)."
 /protein_id="AAH2280.1"
 /db_xref="gi:18490145"
 /translation="MPSIKRLTYITIALCLPSPNMAOCTNFDLDRSGCCLDID
 ECRTPACGDMVCNONGYLCIPRNPYKRPSPYSPYSGPYPAAPLISAP
 NYPTISPLICREYOMDESNOVDSCATDSHOCTQICINTEGGYSCITDGY
 LBGOCADIDECRYCQOLCANYSCTNPNFTLNEDRSCQDNECATENPCV
 OTCVNTYGSFICRCDPELEEDGVHCSDMDECSFSEFLCOHECNOPGTFCSCPC
 YILDNRSQODINECHRNHTCNLOOTCVMLQGFRCIDPICEPEPLRLSDNCKC
 PAENPCRDOPFTLYRDMVVSRSVPADIFOMATTRIFGATYIFQIKSGNEGRF
 YMRQTPISATLVTPRIKPREIQLDLEMTVNTVINFRSSVIRLITYSOYPF"

BASE COUNT 588 a 630 c 547 g 582 t
 ORIGIN

Query Match 100.0%; Score 1269; DB 9; Length 2367;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CAGTGCACGAATGGCTTTGACCTGGATGCGCAGTCAAGAGAGAGTGTATGATATGATGAA 60
 259 CAGTGCACGAATGGCTTTGACCTGGATGCGCAGTCAAGAGAGAGTGTATGATATGATGAA 318
 61 TGGCGAACCATCCCGAGGCGCTGCGGAGAGACATGATGTGTGTAAACCAAAATGGCGGG 120
 319 TGGCGAACCATCCCGAGGCGCTGCGGAGAGACATGATGTGTGTAAACCAAAATGGCGGG 378
 121 TATTATGCAATCCCGGACAAACCTGTGTATCGAGGGCCCTACTCGAACCCCTACTCG 180
 379 TATTATGCAATCCCGGACAAACCTGTGTATCGAGGGCCCTACTCGAACCCCTACTCG 438
 181 ACCCCCTACTGAGTCCGCTTATGCTCCAGAGAGTCCGCTCCAGTCTGCAACTATCC 240
 439 ACCCCCTACTGAGTCCGCTTATGCTCCAGAGAGTCCGCTCCAGTCTGCAACTATCC 498
 241 AGCATCTCCAGGCTCTTATATGCCCTTGTGATACAGATGATGATGAAGAACCAATGT 300
 499 AGCATCTCCAGGCTCTTATATGCCCTTGTGATACAGATGATGATGAAGAACCAATGT 558
 301 GTGGATGTGAGAGAGTGTGACACAGATTTCCACAGTGAACCCACCAAGATCTGCATC 360
 559 GTGGATGTGAGAGAGTGTGACACAGATTTCCACAGTGAACCCACCAAGATCTGCATC 618
 361 AATACTGAAGGCGGATGACCTGCTCTGACACGAGGATATGGCTTCGGAAGGCGAG 420
 619 AATACTGAAGGCGGATGACCTGCTCTGACACGAGGATATGGCTTCGGAAGGCGAG 678
 421 TGGTGAAGATGATGATGATGCTGCTATGCTTACTGACAGAGCTGTGGAATGTCTCT 480
 679 TGGTGAAGATGATGATGCTGCTATGCTTACTGACAGAGCTGTGGAATGTCTCT 738
 481 GGATCTCTTCTTGTGATGATGATGCTTACTGACAGAGCTGTGGAATGTCTCT 540
 739 GGATCTCTTCTTGTGATGATGATGCTTACTGACAGAGCTGTGGAATGTCTCT 798
 541 CAAGATGTGAAGAGTGTGACACGAGACCCCTGCTGCAAAACCTGCGTCAACACTTAC 600
 799 CAAGATGTGAAGAGTGTGACACGAGACCCCTGCTGCAAAACCTGCGTCAACACTTAC 858
 601 GGCCTTTTCATCTGCGCTGTGACCCAGGATATGAACCTTGAGAGAAATGGCGTTTATGC 660
 859 GGCCTTTTCATCTGCGCTGTGACCCAGGATATGAACCTTGAGAGAAATGGCGTTTATGC 918

QY 661 AGTATATGAGACAGATGACCTTCTGAGTTCCTCTGCGCAACATGATGTGTGACACAG 720
 DB 919 AGTATATGAGACAGATGACCTTCTGAGTTCCTCTGCGCAACATGATGTGTGACACAG 978
 QY 721 CCCGGACATATCTTCTGCTCTGCGCTCCGCTCAGGCTACATCTGCTGATGACACAGAGC 780
 DB 979 CCCGGACATATCTTCTGCTCTGCGCTCCGCTCAGGCTACATCTGCTGATGACACAGAGC 1038
 QY 781 TGGCAAGACATCAACGATGTGACAGACAGGACACACAGGACACACAGTGTG 840
 DB 1039 TGGCAAGACATCAACGATGTGACAGACAGGACACACAGGACACACAGTGTG 1098
 QY 841 TACAATTTACAGAGGCGCTTCAATATGATGATGACACACAGGACACACAGTGTG 900
 DB 1099 TACAATTTACAGAGGCGCTTCAATATGATGATGACACACAGGACACACAGTGTG 1158
 QY 901 AGGATGAGATGATTAACCGCTGTATGTCTGCTGAGAACACCTGCTGACAGACACAGC 960
 DB 1159 AGGATGAGATGATTAACCGCTGTATGTCTGCTGAGAACACCTGCTGACAGACACAGC 1218
 QY 961 TTTACCATCTTGTACCGGACATGACAGTGTGTGACAGAGCTCCGCTGAGACATC 1020
 DB 1219 TTTACCATCTTGTACCGGACATGACAGTGTGTGACAGAGCTCCGCTGAGACATC 1278
 QY 1021 TTCCAAATGCAAGCCACGACCGCTTACCTGCGGCTTATTAATTTTCCAGATCAATCT 1080
 DB 1279 TTCCAAATGCAAGCCACGACCGCTTACCTGCGGCTTATTAATTTTCCAGATCAATCT 1338
 QY 1081 GGGATGAGAGGACAGAAATTTTACATGCGGCAAAAGGCGCCCAATGCGCACCCCTGTG 1140
 DB 1339 GGGATGAGAGGACAGAAATTTTACATGCGGCAAAAGGCGCCCAATGCGCACCCCTGTG 1398
 QY 1141 ATGACAGCGCCCATCAAAAGGCGCCGGAATTCACAGCTGAGCTTGGAAATGATCAGTGC 1200
 DB 1399 ATGACAGCGCCCATCAAAAGGCGCCGGAATTCACAGCTGAGCTTGGAAATGATCAGTGC 1458
 QY 1201 AACAGTGCATCACTTGCAGAGAGCTCCGCTGATCCGAGCTGGGATATGTGTGCGAG 1260
 DB 1459 AACAGTGCATCACTTGCAGAGAGCTCCGCTGATCCGAGCTGGGATATGTGTGCGAG 1518
 QY 1261 TACCCTATTC 1269
 DB 1519 TACCCTATTC 1527

RESULT 5

AR036548

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

BASE COUNT

ORIGIN

Query Match 100.0%; Score 1269; DB 6; Length 2550;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CAGTGCACGAATGGCTTTGACCTGGATGCGCAGTCAAGAGAGTGTATGATATGATGAA 60
 457 CAGTGCACGAATGGCTTTGACCTGGATGCGCAGTCAAGAGAGTGTATGATATGATGAA 516


```

crest cells, balloon injured vessels, atherosclerotic
lesions"
/codon.start-1
/product="developmental arteries and neural crest EGF-like
protein"
/protein_id="A4D1768.1"
/db_xref="GI:5305673"
/translation="MPGIRKRLITVTLALCLPSPGNAOACTNGFDLDROSGOCLDID
EERTIPACRGMWVONONGYLCIPRTNIVYRGPSPNPTCIPTMBEGYRCPAPASAP
NPTLSRLICRGYOMDSNOCVDNDECTDHSOCNPTOCITMBEGYRCPAPASAP
LLEGGCLDIDECRYGICQOLCANVPSTCTCNPFITLNEGRSCODNECATENPCV
QTCVNTYSGFICRDPGELEEDGHCSDDECSFSELCQHECVNPGTYFSCPPG
YLLDNRSCDPTINCEBHRNHNCLNLOTCNLOAGFKCIDIRCEBEPRLISDNMCC
PAENPCRDQPFILYRDMDVSGRSVPADIFOMATRYPGAYIIFQIKSGNEREF
YKROTEPISATLIVTRPIKGPRIOLDLEMITVTVINFRSSVIRLRIYVSQPF"
BASE COUNT      560 a      646 c      541 g      581 t
ORIGIN

Query Match      100.0%; Score 1269; DB 9; Length 2328;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CAGTGCACGAATGGCTTTGACCTGGATGCCAGTCAGACAGTGTATGATATGATGAA 60
|||||
244 CAGTGCACGAATGGCTTTGACCTGGATGCCAGTCAGACAGTGTATGATATGATGAA 303
|||||
61 TGGCCGAACATCCCGAGGCGCTGCGAGAGACATGATGTGTAAACCAATGGCGGG 120
|||||
304 TGGCCGAACATCCCGAGGCGCTGCGAGAGACATGATGTGTAAACCAATGGCGGG 363
|||||
121 TATTTATGATTCGCCGAGACAAACCTGTGTATGAGAGGCGCTTACTGACACCCCTACTCG 180
|||||
364 TATTTATGATTCGCCGAGACAAACCTGTGTATGAGAGGCGCTTACTGACACCCCTACTCG 423
|||||
181 ACCCCCTACTCAGTCCGCTACCCAGCAGCTGCCCCACACCTCTGACCTCAACATATCCC 240
|||||
424 ACCCCCTACTCAGTCCGCTACCCAGCAGCTGCCCCACACCTCTGACCTCAACATATCCC 483
|||||
241 AGCATCTCCAGGCGCTTATATGCGCGTTTGATACCATGATGATGAAAGCAACCAATGT 300
|||||
484 AGCATCTCCAGGCGCTTATATGCGCGTTTGATACCATGATGATGAAAGCAACCAATGT 543
|||||
301 GTGGATGTGAGAGAGTGTGCAACAGATTCACACAGTGCACACCCACCAAGATCTGCATC 360
|||||
544 GTGGATGTGAGAGAGTGTGCAACAGATTCACACAGTGCACACCCACCAAGATCTGCATC 603
|||||
361 AATPACGAAGGGGGTACACCTGCTCTGACACGACGATTTGGCTTCTGGAAGGCCAG 420
|||||
604 AATPACGAAGGGGGTACACCTGCTCTGACACGACGATTTGGCTTCTGGAAGGCCAG 663
|||||
421 TGCCTTAGACATGATGAATGCTCCCTAGTGTACTGACGACAGCTCTGTGCGCAATGTCTCT 480
|||||
664 TGCCTTAGACATGATGAATGCTCCCTAGTGTACTGACGACAGCTCTGTGCGCAATGTCTCT 723
|||||
481 GGATCCTATTTCTGTACATGCAACCCCTGTTTTACCTCAATGAGATGAGAGTCTTTC 540
|||||
724 GGATCCTATTTCTGTACATGCAACCCCTGTTTTACCTCAATGAGATGAGAGTCTTTC 783
|||||
541 CAAGATGGAAGAGTGTGCGACGACGACCCCTGCGTGCACAACTGGGTCAACACCTAC 600
|||||
784 CAAGATGGAAGAGTGTGCGACGACGACCCCTGCGTGCACAACTGGGTCAACACCTAC 843
|||||
601 GGCTCTTTCATCTGCGCTGTGACACCGAGATGATGAATGAGAGATGGCTTATTCG 660
|||||
844 GGCTCTTTCATCTGCGCTGTGACACCGAGATGATGAATGAGAGATGGCTTATTCG 903
|||||
661 AGTGAATGAGAGAGTGCAGCTTCTGAGCTTCTGCGCAACATGAGTGTGTAACAG 720
|||||
904 AGTGAATGAGAGAGTGCAGCTTCTGAGCTTCTGCGCAACATGAGTGTGTAACAG 963
|||||
721 CCGGGACATPACTTTCGCTCCTGCGTCCAGGCTACATCTGTGTGATGAGACCAAGC 780
|||||
964 CCGGGACATPACTTTCGCTCCTGCGTCCAGGCTACATCTGTGTGATGAGACCAAGC 1023

```

```

OY 781 TGGCAAGACATCAACGAATGTGAGACAGAGAACCAACAGCTGCAACCTGCAGACAGTGC 840
|||||
DB 1024 TGGCAAGACATCAACGAATGTGAGACAGAGAACCAACAGCTGCAACCTGCAGACAGTGC 1083
|||||
OY 841 TACAATTTACAAAGGGGGCTTCAATGATGATGACCCCATCCGCTGTGAGAGGCTTATCTG 900
|||||
DB 1084 TACAATTTACAAAGGGGGCTTCAATGATGATGACCCCATCCGCTGTGAGAGGCTTATCTG 1143
|||||
OY 901 AGGATCAGTATTAACCCCTGTATGTGTCTCTGCTGAGAACCCCTGGTGCAGAGACCAAGCCC 960
|||||
DB 1144 AGGATCAGTATTAACCCCTGTATGTGTCTCTGCTGAGAACCCCTGGTGCAGAGACCAAGCCC 1203
|||||
OY 961 TTTACCATCTTGAACGGGACATGAGAGTGTGTGACAGACCTCGTCCGCTGACATATC 1020
|||||
DB 1204 TTTACCATCTTGTACGGGGACATGAGAGTGTGTGACAGACCTCGTCCGCTGACATATC 1263
|||||
OY 1021 TTTCAAAATGCAAGCCAGACCCCGCTACCTGCGGCTTATTAATTTCCAGATCAAAATCT 1080
|||||
DB 1264 TTTCAAAATGCAAGCCAGACCCCGCTACCTGCGGCTTATTAATTTCCAGATCAAAATCT 1323
|||||
OY 1081 GGGAAATGAGGGGAGAGAAATTTTACATGCGGCAAAAGGGGCCCATATGATGACCTGTG 1140
|||||
DB 1324 GGGAAATGAGGGGAGAGAAATTTTACATGCGGCAAAAGGGGCCCATATGATGACCTGTG 1383
|||||
OY 1141 ATGACACAGCCCATCAAAAGGGGCCCGGAAATTCAGCTGGAATGATGATCACTGTC 1200
|||||
DB 1384 ATGACACAGCCCATCAAAAGGGGCCCGGAAATTCAGCTGGAATGATGATCACTGTC 1443
|||||
OY 1201 AACACTGTCTATCAACTTCAGAGGAGAGTCCGTGATCCGATGCGGATATATGTGTGCGAG 1260
|||||
DB 1444 AACACTGTCTATCAACTTCAGAGGAGAGTCCGTGATCCGATGCGGATATATGTGTGCGAG 1503
|||||
OY 1261 TACCCATTC 1269
|||||
DB 1504 TACCCATTC 1512
|||||

RESULT 4
LOCUS BC022280
DEFINITION Homo sapiens, clone MGC:22412 IMAGE:4695953, mRNA, complete cds.
ACCESSION BC022280
VERSION BC022280.1 GI:18490144
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2367)
Strausberg, R.
Direct Submission
Submitted (01-FEB-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcdex@lcl.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAL Plate: 36 Row: n Column: 23
This clone was selected for full length sequencing because it

```


	/protein_id="CAB38568.1"	
	/db_xref="GI:4490530"	
	/translation="MPGIRLLIVTITALCLPSPQNAQAQCTNEDFDLDRSGQCLDITIECRITLPEACRGDDMMCVNONGGYLCIPRNPVYRQPSNYPSTPGYPAAPAPLSTANYPNIPISRLICRFQMDENSCQCYDVEINACVADSCDSHCNCTQICINTEGGYTCSDGVYLLSEQCCLDIDECRGYCOQLCANVPGSYSCNCGNFTIEHGRSCQVNECATNPCCQTCVNTGYSFICRGDDPGYELFEEHVGSDMDECSFSEFLCHEVYNQDGTTCSCPYVILLIDNRSQDINIECHERNITCNLOOTCYVLOGGFKIDIRCEEPRLRTSDNRQCMQPAENPGRCHDQFTLLYLRDMDVVSRSVPADIFQWQATTRYGAYITQIKSGNGRREIYMRQTPISATLVMTIRPKGREGRLQDLDEMIVNTVINFRCSSVIRLRIYVSQPF"	
sig_peptide	1..66	
	/gene="FIBL-5"	
mat_peptide	67..1344	
	/gene="FIBL-5"	
	/product="Fibulin-5"	
BASE COUNT	507 a 540 c 475 g 496 t	1 others
ORIGIN		

Query Match	100.0%	Score 1269	DB 9	Length 2019
Best Local Similarity	100.0%	Prod. No. 0		
Matches 1269	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	CAGTGCACGATGGCTTTGACCTGGATCGCCAGTCCAGACAGTGTGTTAGATATTGATGAA	60	
Db	76	CAGTGCACGATGGCTTTGACCTGGATCGCCAGTCCAGACAGTGTGTTAGATATTGATGAA	135	
QY	61	TGCCGAACCAATCCCGGAGGCGCTGCCGAGGAGACATATGTGTGTTAAACCAAAATGGCGGG	120	
Db	136	TGCCGAACCAATCCCGGAGGCGCTGCCGAGGAGACATATGTGTGTTAAACCAAAATGGCGGG	195	
QY	121	TATTTATGCAATTCGCCGAGAAACCCGTGTATGAGAGGCGCTACTGAAACCCCTACTCG	180	
Db	196	TATTTATGCAATTCGCCGAGAAACCCGTGTATGAGAGGCGCTACTGAAACCCCTACTCG	255	
QY	181	ACCCCTCACTCAGSTCCGTACCCAGACAGCTGCCACACTCTCAGCTCCAAACTATACC	240	
Db	256	ACCCCTCACTCAGSTCCGTACCCAGACAGCTGCCACACTCTCAGCTCCAAACTATACC	315	
QY	241	ACGATCTCCAGGCGCTCTTATATGCGCGGTTTGATATACAGATGGATGAAACCAACCAATGT	300	
Db	316	ACGATCTCCAGGCGCTCTTATATGCGCGTGTGGATACCAAGATGATGAAACCAACCAATGT	375	
QY	301	GTCGATGTGGACAGTGTGCACAGATTCCACACAGTGCACACCCACCCAGATCTGCATC	360	
Db	376	GTCGATGTGGACAGTGTGCACAGATTCCACACAGTGCACACCCACCCAGATCTGCATC	435	
QY	361	AATACGGAAGGCGGGTACACCTCTCTCTGCACCGACGGATATTGGCTTCTGGAAGGCCAG	420	
Db	436	AATACGGAAGGCGGGTACACCTCTCTCTGCACCGACGGATATTGGCTTCTGGAAGGCCAG	495	
QY	421	TGCTTAGACATTATGATGAATGCTGCTATGGTTACTGTGCACAGCTGTGTGGAAATGTTCC	480	
Db	496	TGCTTAGACATTATGATGAATGCTGCTATGGTTACTGTGCACAGCTGTGTGGAAATGTTCC	555	
QY	481	GGATCTTATTTCTTGATACATGCAACCTGGTTTTACCCCTCAATGAGATGGAAGCTTTGC	540	
Db	556	GGATCTTATTTCTTGATACATGCAACCTGGTTTTACCCCTCAATGAGATGGAAGCTTTGC	615	
QY	541	CAAGATGTAACGAGTGTGCACCGAGAACCCCTGGTGTGCAAACTGCTCAACACTTAC	600	
Db	616	CAAGATGTAACGAGTGTGCACCGAGAACCCCTGGTGTGCAAACTGCTCAACACTTAC	675	
QY	601	GGCTCTTTCATCTGCGCGTGTACCCAGAGATATGAACCTGAGGAAGATGGCGTTACTTAC	660	
Db	676	GGCTCTTTCATCTGCGCGTGTACCCAGAGATATGAACCTGAGGAAGATGGCGTTACTTAC	735	
QY	661	AATGATATGACAGATGCAAGCTTCTCTGTGAATTCCTGTGCCAAACATGATGTGTGAACCA	720	
Db	736	AATGATATGACAGATGCAAGCTTCTCTGTGAATTCCTGTGCCAAACATGATGTGTGAACCA	795	
QY	721	CCCGGACATATCTTGGTGCCTGGCCCTCCAGGCTACATCTCTGTGATACAAACGGAAC	780	
Db	796	CCCGGACATATCTTGGTGCCTGGCCCTCCAGGCTACATCTCTGTGATACAAACGGAAC	855	

QY	781	TGCCAAGACATCAACGAATGTGAGCAGCAGGAACCCACAGCTGCAACCTGACAGCAGACGTCC	840
Db	886	TGCCAAGACATCAACGAATGTGAGCAGCAGGAACCCACAGCTGCAACCTGACAGCAGACGTCC	915
QY	841	TACAAATTTACAAGGGGGGCTTCAAATGCATCGACCCCATCCGCTGTGAGGAGGCTTATCTG	900
Db	916	TACAAATTTACAAGGGGGGCTTCAAATGCATCGACCCCATCCGCTGTGAGGAGGCTTATCTG	975
QY	901	AGGATCAGTGAATACCGCTGTATGTGTCTCTGCTGAGAAACCTTGGCTGCAGAACACACGCC	960
Db	976	AGGATCAGTGAATACCGCTGTATGTGTCTCTGCTGAGAAACCTTGGCTGCAGAACACACGCC	1033
QY	961	TTTACCATCTTGTATACCGGGAGCATGAGCCTGTGTTGAGAGACGCTCCGTTCCCGCTACATC	1020
Db	1036	TTTACCATCTTGTATACCGGGAGCATGAGCCTGTGTTGAGAGACGCTCCGTTCCCGCTACATC	1099
QY	1021	TTTCCAAATGCAAGCCACGACCCCGCTACCTTGGGGCTTATTCATTTTCCAGATCAAAATCT	1086
Db	1096	TTTCCAAATGCAAGCCACGACCCCGCTACCTTGGGGCTTATTCATTTTCCAGATCAAAATCT	1155
QY	1081	GGGATGAGGGCAGAGAAATTTTACATGCGGGCAAAAGGGGCCCATAGTGGCCACCCTGGGTG	1144
Db	1156	GGGATGAGGGCAGAGAAATTTTACATGCGGGCAAAAGGGGCCCATAGTGGCCACCCTGGGTG	1211
QY	1141	ATGACAGCGCCCATCAAAAGGGGCCCGGGGAAATTCACGTGCACTTGGAAATGATCACTGTG	1200
Db	1216	ATGACAGCGCCCATCAAAAGGGGCCCGGGGAAATTCACGTGCACTTGGAAATGATCACTGTG	1277
QY	1201	AACACTGTCACTCACTTCAGAGGAGGCTCCGTGATCCGATCGCGGATTTATGTGTGCGGAG	1266
Db	1276	AACACTGTCACTCACTTCAGAGGAGGCTCCGTGATCCGATCGCGGATTTATGTGTGCGGAG	1335
QY	1261	TACCCATTTC 1269	
Db	1336	TACCCATTTC 1344	

[illegible]

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="14"
/map="14q32.1"
169 1515
/notes="DANCE; contains 6 CBECF domains and 1 RGD motif;
secreted protein; expressed in developmental aorta, neural

```

FEATURES
source

CDS

Location/Qualifiers
1..1720
/organism="unidentified"
/db_xref="taxon:32644"
211..1157
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB69285.1"
/db_xref="GI:6733225"
/translation="MRKIRILVTLLALCLPSRGNACQCTNFPDIDROSGCLDID
ECRTIPACRDMCVNONGYICLIPRTNVRGYSNPSTIPSTSGPTPAAPLISAP
NPTISRLPCHRYQMDENOCVDDECATSHQCNPTICIMTEGGYTCSDIGYM
LLEGCLDIDECRYGCOOLCANVPISCTCNPGFTLNEDRSODVNECATENPCY
QYCVATYGSFICRDPYLEDYSDSDDECSEFLCOHEVCNPGTYGSCPPG
YLLDNRSCODINCEHNRHNTNLOOTCYNLOGFKCIPDIRCEEPYLRISDRNCM
PAENGCRRDPFTILYRDMDVYSRSPALIFOMQATRRYRGATYIRIOIKSGNREF
YMRQIRISATILVMTPIKPIRPIQDLDEMTVNTVINFRRSSVIRLRITYSIRP"
sig_peptide 211..284
mat_peptide 287..1354
BASE COUNT 396 a 515 c 434 g 372 t 3 others
ORIGIN

Query Match 100.0%; Score 1269; DB 6; Length 1720;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CAGTCACGATGGCTTTGACCTGATGCGCAGTCAGACAGTGTATGATTTGATGATAA 60
286 CAGTCACGATGGCTTTGACCTGATGCGCAGTCAGACAGTGTATGATTTGATGATAA 345
61 TGGCGAACCATCCCGAGGCGCTGCGAGAGACATGATGTGTAAACAAATGGCGGG 120
346 TGGCGAACCATCCCGAGGCGCTGCGAGAGACATGATGTGTAAACAAATGGCGGG 405
121 TATTTATGATTCCTCCCGAGAACACCTGTGTATGAGAGGCCCTTACTGCAACCCCTACTCG 180
406 TATTTATGATTCCTCCCGAGAACACCTGTGTATGAGAGGCCCTTACTGCAACCCCTACTCG 465
181 ACCCCCTACTGAGTCCGTACCCAGCAGTGCACCACTCTCAGCTCCAACTATCC 240
466 ACCCCCTACTGAGTCCGTACCCAGCAGTGCACCACTCTCAGCTCCAACTATCC 525
241 ACGATCTCCAGGCGCTTATATGCGCGTTTGATACCAAGATGATGAAGCAACCAATGT 300
526 ACGATCTCCAGGCGCTTATATGCGCGTTTGATACCAAGATGATGAAGCAACCAATGT 585
301 GTGGATGTGGAGAGTGTGCAACAGATTCACCAGAGTGCACACCCCACTATCTGCATC 360
586 GTGGATGTGGAGAGTGTGCAACAGATTCACCAGAGTGCACACCCCACTATCTGCATC 645
361 AATATCTGAAGGGGGGTACACCTCTCTGACCGAGGATATTGGCTTGTGAAGGCCAG 420
646 AATATCTGAAGGGGGGTACACCTCTCTGACCGAGGATATTGGCTTGTGAAGGCCAG 705
421 TGCCTAGACATTTGATGATGTGCTATGTTACTGCGCAGCAGCTGTGCGCAATGTTTCT 480
706 TGCCTAGACATTTGATGATGTGCTATGTTACTGCGCAGCAGCTGTGCGCAATGTTTCT 765
481 GGATCCTATCTTGTTCATGACGCAACCTGGTTTACCTCAATGAGATGGAAGTCTTGC 540
766 GGATCCTATCTTGTTCATGACGCAACCTGGTTTACCTCAATGAGATGGAAGTCTTGC 825
541 CAAGATGTGAAGAGTGTGCGACCGAGAACCCCTGCTGGAACCTGCTGTAACACCTTAC 600
826 CAAGATGTGAAGAGTGTGCGACCGAGAACCCCTGCTGGAACCTGCTGTAACACCTTAC 885
601 GGCTCTTTTCATCTGCCGTGTGACCGAGATATGAACCTTGAAGAAGATGGCTTCATTGC 660
886 GGCTCTTTTCATCTGCCGTGTGACCGAGATATGAACCTTGAAGAAGATGGCTTCATTGC 945
661 AGTGATATGAGAGAGTGTGCTGATGTTCTCTGACCAACATGATGTGTGAACAG 720
946 AGTGATATGAGAGAGTGTGCTGATGTTCTCTGACCAACATGATGTGTGAACAG 1005

721 CCCGACATACCTTCTGCTCTGCTGCTCCAGGCTACATCTGTGTATGACACCGAAGC 780
1006 CCCGACATACCTTCTGCTCTGCTGCTCCAGGCTACATCTGTGTATGACACCGAAGC 1065
781 TGGCAAGACATCAACGAATGTGAGCAGACGAGAACCAACAGTGCACCTGACAGACGTGC 840
1066 TGGCAAGACATCAACGAATGTGAGCAGACGAGAACCAACAGTGCACCTGACAGACGTGC 1125
841 TACAAATTTACAAAGGGGCTTCAATGATGATGACCCCATCCCTGTGAGAGAGCTTATCTG 900
1126 TACAAATTTACAAAGGGGCTTCAATGATGATGACCCCATCCCTGTGAGAGAGCTTATCTG 1185
901 AGGATGATGATTAACCGCTGATGTGCTGCTGCTGAGAACCTGTGCTGAGACAGACGCC 960
1186 AGGATGATGATTAACCGCTGATGTGCTGCTGCTGAGAACCTGTGCTGAGACAGACGCC 1245
961 TTTACCAATCTTGTACCGGGACATGAGAGTGTGTCAGAGCGCTCCGTTCCGCTGACATC 1020
1246 TTTACCAATCTTGTACCGGGACATGAGAGTGTGTCAGAGCGCTCCGTTCCGCTGACATC 1305
1021 TTTCAATGTACAGCCACGACCCGCTACCTGCGGCGCTATTACATTTTCCAGATCAAAATCT 1080
1306 TTTCAATGTACAGCCACGACCCGCTACCTGCGGCGCTATTACATTTTCCAGATCAAAATCT 1365
1081 GGGATGAGAGGAGAGAAATTTTACATGCGGCAACGGGCGCCCATCAGTGCACCCCTGTGTG 1140
1366 GGGATGAGAGGAGAGAAATTTTACATGCGGCAACGGGCGCCCATCAGTGCACCCCTGTGTG 1425
1141 ATGACACGCCCCATCAAAAGGGCCCCGGGAAATTCAGCTGAGCTTGAATGATCACTGTG 1200
1426 ATGACACGCCCCATCAAAAGGGCCCCGGGAAATTCAGCTGAGCTTGAATGATCACTGTG 1485
1201 AACACTGTCACTCACTCAAGAGAGCGTCCGTAATCCAGTGGGATATATGTGTGCGAG 1260
1486 AACACTGTCACTCACTCAAGAGAGCGTCCGTAATCCAGTGGGATATATGTGTGCGAG 1545
1261 TACCCTATTC 1269
1546 TACCCTATTC 1554

RESULT 2
HSA133490
LOCUS HSA133490 2019 bp mRNA linear PRI 11-MAR-1999
DEFINITION Homo sapiens fibulin-5.
ACCESSION AJ133490
VERSION AJ133490.1 GI:4490529
KEYWORDS FIBL-5 gene; fibulin-5.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2019)
AUTHORS Kostka G.
TITLE Direct Submission
JOURNAL Submitted (05-MAR-1999) Kostka G., Dept. of Protein Chemistry,
Max-Planck-Institut fuer Biochemie, Am Klopferspitz, 82152
Martinsried, GERMANY
REFERENCE 2 (bases 1 to 2019)
AUTHORS Kostka G.
JOURNAL Unpublished
FEATURES
source 1..2019
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/issue_type="melanoma"
1..1347
/gene="FIBL-5"
1..1347
/gene="FIBL-5"
/codon_start=1
/product="fibulin-5"

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 08:43:26 ; Search time 3309.01 Seconds

(Without alignments)
11160.894 Million cell updates/sec

File: US-09-674-379A-15

Sequence: 1 cagtgacgaatgacttga.....atgttcgcagtaaccattc 1269

Scoring table: IDENTITY_NIC
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: GenBank
1: gb|000001.1
2: gb|000002.1
3: gb|000003.1
4: gb|000004.1
5: gb|000005.1
6: gb|000006.1
7: gb|000007.1
8: gb|000008.1
9: gb|000009.1
10: gb|000010.1
11: gb|000011.1
12: gb|000012.1
13: gb|000013.1
14: gb|000014.1
15: gb|000015.1
16: gb|000016.1
17: gb|000017.1
18: gb|000018.1
19: gb|000019.1
20: gb|000020.1
21: gb|000021.1
22: gb|000022.1
23: gb|000023.1
24: gb|000024.1
25: gb|000025.1
26: gb|000026.1
27: gb|000027.1
28: gb|000028.1
29: gb|000029.1
30: gb|000030.1
31: gb|000031.1
32: gb|000032.1
33: gb|000033.1
34: gb|000034.1
35: gb|000035.1
36: gb|000036.1
37: gb|000037.1
38: gb|000038.1
39: gb|000039.1
40: gb|000040.1
41: gb|000041.1

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1269	100.0	1720	6	A84086 Sequence 1
2	1269	100.0	2328	9	AJ133490 Homo sapi
3	1269	100.0	2367	9	AF112152 Homo sapi
4	1269	100.0	2550	6	BC022280 Homo sapi
5	1269	100.0	2550	6	AR036548 Homo sapi
6	1269	100.0	2550	6	AR173204 Homo sapi
7	1267.4	99.9	2609	6	AX376340 Sequence
8	1267.4	99.9	2609	6	AX403659 Sequence
9	1253	98.7	2019	9	AF093118 Homo sapi
10	1056.2	83.2	2230	10	BC006636 Mus muscu
11	1056.2	83.2	2478	10	AF112151 Homo sapi
12	1046.6	82.5	2499	10	AF112153 Homo sapi
13	1045	82.3	2304	10	AF137350 Homo sapi
14	372.8	29.4	1536	9	HSA133819 Homo sapi
15	372.8	29.4	1536	9	AB030655 Homo sapi
16	371.2	29.3	1480	6	AX023967 Sequence
17	371.2	29.3	1561	9	BC010456 Homo sapi
18	371.2	29.3	1707	9	AK000980 Homo sapi
19	371.2	29.3	1875	6	AX201326 Homo sapi
20	371.2	29.3	1875	6	AX464184 Homo sapi
21	371.2	29.3	2018	6	AR095382 Homo sapi
22	369.6	29.1	1757	9	AF109121 Homo sapi
23	364.8	28.7	1825	9	AF124486 Homo sapi
24	363.8	28.7	1705	9	AF093119 Homo sapi
25	350.4	27.6	1778	10	AF046870 Homo sapi
26	347.2	27.4	1513	10	AF104223 Homo sapi
27	347.2	27.4	1513	10	AX023961 Homo sapi
28	347.2	27.4	1794	10	AF109122 Homo sapi
29	345.6	27.2	1724	10	BC012265 Homo sapi
30	337.6	26.6	1358	6	AX023976 Sequence
31	312	24.6	1021	6	AX023954 Sequence
32	295.8	23.3	2613	9	AK098186 Homo sapi
33	268	21.1	1122	6	AX023965 Sequence
34	260.6	20.5	171901	9	AC007057 Homo sapi
35	260.6	20.5	172674	9	CNS00008 Homo sapi
36	260.6	20.5	175210	9	CNS07EG5 Homo sapi
37	250.6	19.7	1677	10	BC031184 Homo sapi
38	249	19.6	1911	10	D89730 Homo sapi
39	241	19.0	2025	9	BC014410 Homo sapi
40	241	19.0	2512	6	AX277602 Sequence
41	241	19.0	2512	6	AX331916 Sequence
42	241	19.0	2512	6	AX332783 Sequence
43	241	19.0	2512	6	AX333462 Sequence
44	241	19.0	2512	6	AX334103 Sequence
45	241	19.0	2512	6	AX335899 Sequence

ALIGNMENTS

RESULT 1
LOCUS A84086 1720 bp DNA
DEFINITION Sequence 1 from Patent WO9846746.
ACCESSION A84086
VERSION A84086.1 GI:6733224
KEYWORDS
SOURCE unclassified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1720)
AUTHORS Li, H. and Olsen, H. S.
TITLE EXTRACELLULAR/EPIDERMAL GROWTH FACTOR LIKE PROTEIN
JOURNAL Patent: WO 9846746-A 1 22-OCT-1998;
HUMAN GENOME SCIENCES INC (US); LI HAODONG (US)

QY 1021 TTCGAATGCAAGCCAGACCCGCTACCTGGGGCTATTACATTTTCAGATCAAAATCT 1080
1096 TTCAGATGCAAGCCAGACCCGCTACCTGGGGCTATTACATTTTCAGATCAAAATCT 1155
QY 1081 GGGAAATGAGGGGAGAGAAATTTTACATGCGGCAAAAGGGCCCATCAGTCCACCCCTGGTG 1140
1156 GGCACAGAGGGTGCAGAGTCTATATGCGGCAAAAGGGCCCATCAGTCCACCCCTGGTG 1215
QY 1141 ATGACAGCGCCCATCAAGGGCCCGGAAATCCAGCTGACCTGGAATGATCACTGTC 1200
1216 ATGACAGCGCCCATCAAGGGCCCGGAAATCCAGCTGACCTGGAATGATCACTGTC 1275
QY 1201 AACACTGTATCAACTTTCAGAGGAGCTCCGTGATCCGACTGCGGATATATGTGCGCAG 1260
1276 AACACTGTATCAACTTTCAGAGGAGCTCCGTGATCCGACTGCGGATATATGTGCGCAG 1335
QY 1261 TACCCATTC 1269
Db 1336 TATCCGTTT 1344

Search completed: July 3, 2003, 11:59:31
Job time : 283.77 secs

Db 1022 TCTGGGAATGAGGGGAGAGATTTTTCATGCGGCAAAAGGGCCCATTCATGCGACCCCTG 1081
 QY 1138 GTGATGACACGCCCCATCAAGAGGCCCGGGAATACAGTGGACTTGGAAATGATGACT 1197
 Db 1082 GTGATGACACGCCCCATCAAGAGGCCCGGGAATACAGTGGACTTGGAAATGATGACT 1141
 QY 1198 GTCAACACGTGATCACTTCAAGAGGACCTCGTATCCGACTGCGGATATATGTGTCG 1257
 Db 1142 GTCAACACGTGATCACTTCAAGAGGACCTCGTATCCGACTGCGGATATATGTGTCG 1201
 QY 1258 CAGTACCCATTC 1269
 Db 1202 CAGTACCCATTC 1213

RESULT 13
 AA239385
 AA239385 standard; DNA; 1269 BP.

AA239385;
 22-FEB-2000 (first entry)

Smooth muscle proliferation modulating mature protein coding sequence.

Cell proliferation; vascular; smooth muscle cell; arteriosclerosis; PTCA; endothelial thickening; percutaneous transluminal coronary angioplasty; myoma; hematopoietic cell-regulation; cytokine; tissue generation; taxis; actin activity; chemotaxis; blood coagulation; thrombotic; tumor; mouse; metastasis; nutrient; ss.

Mus musculus.

MO9955863-A1.

04-NOV-1999.

28-APR-1999; 99MO-JP02283.

28-APR-1998; 98JP-0119731.

(ONOX) ONO PHARM CO LTD.

Honjo T, Tashiro K, Nakamura T;

WPI: 2000-038646/03.

P-PSDB; AAY56751.

Novel mouse polypeptides for treatment, e.g. arteriosclerosis and myoma

Claim 4; Page 53-54; 70pp; Japanese.

The invention provides mouse polypeptides for treatment of diseases due to abnormal proliferation of smooth muscle. The polypeptides can be produced by standard recombinant methodology. The polypeptides can be used according to their inhibition of the proliferation of vascular smooth muscle cells, particularly in treating arteriosclerosis or re-transluminar coronary angioplasty (PTCA), or myoma, hematopoietic cell-regulatory activity, cytokine activity, tissue generation/repair, blood coagulation/thrombotic activity, receptor/ligand activity, cadherin/actin activity, tumor metastasis inhibiting activity, tumor inhibition, and as nutrient. The present sequence represents the coding sequence of the mature protein cell proliferation.

Sequence 1269 BP; 304 A; 350 C; 314 G; 301 T; 0 other;

Query Match 83.2%; Score 1056.2; DB 21; Length 1269;
 Best Local Similarity 89.5%; Pred. No. 0;

	Matches 1136; Conservative 0; Mismatches 133; Indels 0; Gaps 0;
QY 1	CAGTGCAGAAATGGCTTTGACCTGGATGCGCAGTACAGACAGTGTATGATATGATGAA 60
Db 1	CAGTGCAGAAATGGCTTTGACCTGGATGCGCAGTACAGACAGTGTATGATATGATGAA 60
QY 61	TGCGGAACATCCCGGAGGCTGGGAGAGACATGATGTGTAAACAAATGGCGGG 120
Db 61	TGCGGAACATCCCGGAGGCTGGGAGAGACATGATGTGTAAACAAATGGCGGG 120
QY 121	TATTTATGATTCCTCCGAGCAAAACCTGTATGAGAGGCTTACTGAAACCTTACTCG 180
Db 121	TATTTATGATTCCTCCGAGCAAAACCTGTATGAGAGGCTTACTGAAACCTTACTCG 180
QY 181	ACCCCTACTAGAGTCCGATACCAAGACAGTCCGACACTGACCTGCAACATATCC 240
Db 181	ACATCTCTACTAGAGTCCGATACCAAGACAGTCCGACACTGACCTGCAACATATCC 240
QY 241	ACGATCTCCAGGCTCTTATATGCGCTTGTGATACCAAGATGATGAAACCAATAT 300
Db 241	ACGATCTCCAGGCTCTTATATGCGCTTGTGATACCAAGATGATGAAACCAATAT 300
QY 301	GTGATGTGAGAGTGTGCAACAGATTCACCAAGTGCACCCACCAAGATCTGATAT 360
Db 301	GTGATGTGAGAGTGTGCAACAGATTCACCAAGTGCACCCACCAAGATCTGATAT 360
QY 361	AATACGAAAGGCGGTACACCTGCTCCGACAGAGATATGCTTCTGGAAGCCAG 420
Db 361	AACATGGAAGAGGTATACCTGCTCCGACAGAGATATGCTTCTGGAAGCCAG 420
QY 421	TGCTTAGACATGATGAAATGTGCTATGCTGCTGACAGACCTGTCGAAATGTCT 480
Db 421	TGCTTAGACATGATGAAATGTGCTATGCTGCTGACAGACCTGTCGAAATGTCT 480
QY 481	GGATCCATATCTTGTACATGGAACCTGTTTACCTCATAGAGATGGAAGTCTTGC 540
Db 481	GGATCCATATCTTGTACATGGAACCTGTTTACCTCATAGAGATGGAAGTCTTGC 540
QY 541	CAAGATGGAAGAGTGTGCAACAGATTCCTGCAAAACCTGCTGCAACACCTAT 600
Db 541	CAAGATGGAAGAGTGTGCAACAGATTCCTGCAAAACCTGCTGCAACACCTAT 600
QY 601	GGCTCTTTCATCTGCGCTGTGACCCAGATATGACCTTGAAGAAATGCGCTTATGC 660
Db 601	GGCTCTTTCATCTGCGCTGTGACCCAGATATGACCTTGAAGAAATGCGCTTATGC 660
QY 661	AGTATATGAGAGAGTGTGCAACAGATTCCTGCAAAACCTGCTGCAACACCTAT 720
Db 661	AGTATATGAGAGAGTGTGCAACAGATTCCTGCAAAACCTGCTGCAACACCTAT 720
QY 721	CCGCGCATATCTTCTCTCTGCTGCTGCAAGCTTACCTCTGCTGATGACCAACGAGC 780
Db 721	CCGCGCATATCTTCTCTCTGCTGCTGCAAGCTTACCTCTGCTGATGACCAACGAGC 780
QY 781	TGCCAAGACATCAACAGATGAGACAGAGACAGACAGAGTACCTGCAACCTGCAACGTCG 840
Db 781	TGCCAAGATATCAATGATGAGACAGAGACAGAGTACCTGCAACCTGCAACGTCG 840
QY 841	TACAAATTTACAGAGGCTTCAAAATGATGATGATGATGATGATGATGATGATGATGAT 900
Db 841	TACAAATTTACAGAGGCTTCAAAATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 901	AGGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db 901	AGGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 961	TTTACCATCTTATACCGGAGATGAGAGTGTGCTGAGAGAGCTCCCTCCCGCTGATC 1020
Db 961	TTTACCATCTTATACCGGAGATGAGAGTGTGCTGAGAGAGCTCCCTCCCGCTGATC 1020
QY 1021	TTTACCATCTTATACCGGAGATGAGAGTGTGCTGAGAGAGCTCCCTCCCGCTGATC 1080
Db 1021	TTTACCATCTTATACCGGAGATGAGAGTGTGCTGAGAGAGCTCCCTCCCGCTGATC 1080

Db	1725	AACGCTTCATCACTTCAGAGCAGCTCCCTGATCCGACTGCGGATATATGTGTCCAG	1784
Qy	1261	TACCCATTC	1269
Db	1785	TACCCATTC	1793

RESULT 12
AAT89380
ID AAT89380 standard; cDNA; 1228 BP,
vv

DT 14-APR-1998 (first entry)
yy

Human extracellular growth factor-like protein cDNA.

Kidney disorder; liver disorder; embryogenesis; angiogenesis; therapy; ss.

OS Homo sapiens.

FH	Key	Location/Qualifiers
FT	CDS	38..1216

PN WO9738002-A1.

PD 16-OCT-1997.

PF 10-APR-1996; 96WO-US05247.

PR 10-APR-1996; 96WO-US05247.

PA (HUMA-) HUMAN GENOME SCI INC.

PI L1 H, Olsen HS;

DR WPI; 1997-512640/47.

[illegible]

PT Nucleic acid encoding extracellular/epidermal growth factor-like protein - useful for treatment and diagnosis of e.g. wounds, neurological disease, neoplasia, psoriasis etc

PS Claim 4; Fig 1A-D; 48pp; English.

CC This claimed cDNA clone codes for a human polypeptide (see MAM31705).
CC identified on the basis of homology as a putative extracellular
CC protein-like/epidermal growth factor-like (EGF) protein. It was
CC isolated from a human foetal heart cDNA library, and can be
CC amplified from ATCC 97285 and the product cloned into vectors for
CC expression in bacterial, insect, mammalian or plant cells. EGF
CC polypeptides, and polynucleotides encoding them, can be used e.g.
CC to induce DNA synthesis, to regulate vascular smooth muscle
CC proliferation, to treat Marfan syndrome, to stimulate wound healing,
CC to restore normal neurological function after trauma or AIDS
CC dementia, to treat ocular disorders, to treat kidney and liver
CC disorders, to promote hair follicular development, to stimulate
CC growth and differentiation of epidermal and epithelial cells in
CC vivo and in vitro, for the treatment of burns, ulcers and corneal
CC incisions, and to stimulate embryogenesis and angiogenesis.
CC Nucleic acid fragments are also used as probes or primers, e.g. for
CC diagnosis and chromosome identification.

Sequence 1228 BP; 293 A; 364 C; 300 G; 271 T; 0 other;

Query Match	95.1%;	Score 1207.2;	DB 18;	Length 1228;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 1209;	Conservative	0;	Mismatches 3;	Indels 0;
			Gaps	0;

OY	58	GAATGCCAACCATTCCCGAGGCGCGGAGAGACATGATGTGTTAAACCMAATGGC	117
Db	2	GAGTGCAGCAACCATCCCGAGGCGCTGCCGAGAGACATGATGTGTTAAACCMAATGGC	61
OY	118	GGGTATTTTATGCATTTCCCGGAGCAAAACCCGTGTATTCGAGGGGCCCTACTCGAACCCCTAC	177
Db	62	GGGTATTTTATGCATTTCCCGGAGCAAAACCCGTGTATTCGAGGGGCCCTACTCGAACCCCTAC	121
OY	178	TCGACCCCCCTACTAGGTCCCTACCACACACTGGCCACACACTCTCAGCTTCGAAACTAT	237
Db	122	TCGACCCCCCTACTAGGTCCCTACCACACACTGGCCACACACTCTCAGCTTCGAAACTAT	181
OY	238	CCCAACGATCTCCAGGCGCTTATATATAGCCGCTTTGGATACAGATGGATGAAGCAACCAA	297
Db	182	CCCAACGATCTCCAGGCGCTTATATATAGCCGCTTTGGATACAGATGGATGAAGCAACCAA	241
OY	298	TGTGTGGATGTGACAGAGTGTGCAACAGATTCACACAGTGCACCCACCCAGATCTGC	357
Db	242	TGTGTGGATGTGACAGAGTGTGCAACAGATTCACACAGTGCACCCACCCAGATCTGC	301
OY	358	ATCAATACTGAAGGCGGGTACACCGTCGCTGCACCGAGGATATTTGGCTTTGGAAAGGC	417
Db	302	ATCAATACTGAAGGCGGGTACACCGTCGCTGCACCGAGGATATTTGGCTTTGGAAAGGC	361
OY	418	CAGTGCCTTAGACATTTGATGAATGTGCGCTATGGTACTGCCACAGCAGCTCTGTGCCAGATTT	477
Db	362	CAGTGCCTTAGACATTTGATGAATGTGCGCTATGGTACTGCCACAGCAGCTCTGTGCCAGATTT	421
OY	478	CCTGATTCCTATTTCTGTACATGCACACCCCTGGTTTTACCCTTCATAGAGATGGAAGTCT	537
Db	422	CCTGATTCCTATTTCTGTACATGCACACCCCTGGTTTTACCCTTCATAGAGATGGAAGTCT	481
OY	538	TGCCAAGATGGAAGAGATGTGCCACAGCAACCCCTCGTGTCAAACTGTGCTCAACAC	597
Db	482	TGCCAAGATGGAAGAGATGTGCCACAGCAACCCCTCGTGTCAAACTGTGCTCAACAC	541
OY	598	TAGGAGCTTTTCATCTGCGCTGTGACCCAGAGATATGCACTTGAGGAAGATGGCGTTCAT	657
Db	542	TAGGAGCTTTTCATCTGCGCTGTGACCCAGAGATATGCACTTGAGGAAGATGGCGTTCAT	601
OY	658	TGCAGTGTATGAGAGAGATGTGACGCTTCCTAGATTCCTGTGCAACATAGAGTGTGAAC	717
Db	602	TGCAGTGTATGAGAGAGATGTGACGCTTCCTAGATTCCTGTGCAACATAGAGTGTGAAC	661
OY	718	CAGGCGGCAATATCTTGTGCTGCTGCGCTCCAGAGCTACATCTCTGCTGGATGACAAACGA	777
Db	662	CAGGCGGCAATATCTTGTGCTGCTGCGCTCCAGAGCTACATCTCTGCTGGATGACAAACGA	721
OY	778	AGCTTGCCACAGATCAACGAATGTGAGCACAGGAACCAACAGTGCACACTGTGACGACAG	837
Db	722	AGCTTGCCACAGATCAACGAATGTGAGCACAGGAACCAACAGTGCACACTGTGACGACAG	781
OY	838	TGTACACATTTACAAAGGGGGCTTCGAAATGATCGACCCATCCGCTGTGAGAGAGCTTAT	897
Db	782	TGTACACATTTACAAAGGGGGCTTCGAAATGATCGACCCATCCGCTGTGAGAGAGCTTAT	841
OY	898	CTGAGGATACAGTGAATACCGCTGTATGTCTGTGTGAAACCCCTGCTGCAGAGACCG	957
Db	842	CTGAGGATACAGTGAATACCGCTGTATGTCTGTGTGAAACCCCTGCTGCAGAGACCG	901
OY	958	CCCTTTACCATCTTGTACCGGGACATGAGAGCTGTGTAGAGAGCTCCGCTTCGCCCTGAC	1017
Db	902	CCCTTTACCATCTTGTACCGGGACATGAGAGCTGTGTGTAGAGAGCTCCGCTTCGCCCTGAC	961
OY	1018	ATCTTCCAAATGCAAGGACGACGACCGGCTACCCCTGGGGGCTTATACATTTTCCAGATCAA	1077
Db	962	ATCTTCCAAATGCAAGGACGACGACCGGCTACCCCTGGGGGCTTATACATTTTCCAGATCAA	1021
OY	1078	TCCTGGGAATGAGGGCAGAGATTTTTCATGCGGCAAAACGGGGCCCATCATGTCGACCCCTG	1137


```

Db      765  ACGATCTCCAGGCGCTTATATGCCGCTTGATACCGAGTGAAGCAACCAATGT 824
QY      301  GTGATGTGACGAGTGTGCAACAGATTCCACAGTCACACCCAGATCTGCATC 360
Db      825  GTGATGTGACGAGTGTGCAACAGATTCCACAGTCACACCCAGATCTGCATC 884
QY      361  AATCTGAAGGGGGTACACTGCTCTGACCGAGATATGCTTCTGGAAGCCAG 420
Db      885  AATCTGAAGGGGGTACACTGCTCTGACCGAGATATGCTTCTGGAAGCCAG 944
QY      421  TGGTTAGACATTGTAATGTGCTATGCTTACTGCGACAGCTGTGCGAATGTCCT 480
Db      945  TGGTTAGACATTGTAATGTGCTATGCTTACTGCGACAGCTGTGCGAATGTCCT 1004
QY      481  GGAATCCATTTCTTATACATGCAACCTGCTGTTTAACTCAATGAGATGGAAGTCTTGC 540
Db      1005  GGAATCCATTTCTTATACATGCAACCTGCTGTTTAACTCAATGAGATGGAAGTCTTGC 1064
Y       541  CAAGATGTGAAGAGTGTGCGACCGAGAACCCCGCGGCAACCTGGGTCAACACCTAC 600
Db      1065  CAAGATGTGAAGAGTGTGCGACCGAGAACCCCGCGGCAACCTGGGTCAACACCTAC 1124
QY      601  GGGCTTTTATCTGCGCGCTGTGACCGAGATATGAACTTGAAGAGATGGCTTCATTC 660
Db      1125  GGGCTTTTATCTGCGCGCTGTGACCGAGATATGAACTTGAAGAGATGGCTTCATTC 1184
QY      661  AGTATATGAGAGAGTGCACCTTCTGAGTCTCTGCTGCGCAACATGAGTGTGACAG 720
Db      1185  AGTATATGAGAGAGTGCACCTTCTGAGTCTCTGCTGCGCAACATGAGTGTGACAG 1244
QY      721  CCCGGACATACTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db      1245  CCCGGACATACTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1304
QY      781  TGGCAAGATCATCAGCAATGTGAGCAGACAGAACCAACGTCGCAACCTGTCAGCAG 840
Db      1305  TGGCAAGATCATCAGCAATGTGAGCAGACAGAACCAACGTCGCAACCTGTCAGCAG 1364
QY      841  TACAAATTTACAAGGGGGCTTCAATGATGATGACACCCCTGCTGAGAGAGCCCTTACTG 900
Db      1365  TACAAATTTACAAGGGGGCTTCAATGATGATGACACCCCTGCTGAGAGAGCCCTTACTG 1424
QY      901  AGGATCAGTATACCGCTGTATGTCTGCTGAGAACCCCTGCTGAGAGAGCCCTTACTG 960
Db      1425  AGGATCAGTATACCGCTGTATGTCTGCTGAGAACCCCTGCTGAGAGAGCCCTTACTG 1484
Y       961  TTACCATCTTGTACCGGGAGCAATGAGAGTGTGCTGAGAACCCCTGCTGAGAGAGCC 1020
Db      1485  TTACCATCTTGTACCGGGAGCAATGAGAGTGTGCTGAGAACCCCTGCTGAGAGAGCC 1544
QY      1021  TTCCAATGCAAGCCAGCAGACCGCTACCTGCGGCTTATTACATTTTCCAGATCAATCT 1080
Db      1545  TTCCAATGCAAGCCAGCAGACCGCTACCTGCGGCTTATTACATTTTCCAGATCAATCT 1604
QY      1081  GGGATGAGGGGAGAGAAATTTTACATGCGGCAACCGGCCCAATCAGTGCACCTGCTG 1140
Db      1605  GGGATGAGGGGAGAGAAATTTTACATGCGGCAACCGGCCCAATCAGTGCACCTGCTG 1664
QY      1141  ATGACACGCCCCATCAAGGGGGCCCGGGAAATTCACGCTGAGACTTGAAGATGACACGTC 1200
Db      1665  ATGACACGCCCCATCAAGGGGGCCCGGGAAATTCACGCTGAGACTTGAAGATGACACGTC 1724
QY      1201  AACACTGTCATCACTTCAAGAGGAGCTCGTATCCGATCGAGTGGGATATATGTGTGAG 1260
Db      1725  AACACTGTCATCACTTCAAGAGGAGCTCGTATCCGATCGAGTGGGATATATGTGTGAG 1784
QY      1261  TACCATTC 1269
Db      1785  TACCATTC 1793

```

RESULT 10

```

AAS46128
ID      AAS46128 standard; cDNA; 2609 bp.
XX
AC      AAS46128;
XX
DT      18-DEC-2001 (first entry)
XX
DE      Human DNA encoding PRO polypeptide sequence #204.
XX
KW      dog, cat; pig; goat; rabbit; tumour; cancer; human; cattle; horse; sheep; ss;
KW      blood; chromosome cell; cell proliferation; cell differentiation; colon;
KW      adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder;
XX      PCR primer.
XX
OS      Homo sapiens.
XX
PN      WO200168848-A2.
XX
PD      20-SEP-2001.
XX
PE      28-FEB-2001; 2001WO-US06520.
XX
PR      01-MAR-2000; 2000WO-US05601.
PR      02-MAR-2000; 2000WO-US05841.
PR      03-MAR-2000; 2000US-187202P.
PR      06-MAR-2000; 2000US-186968P.
PR      14-MAR-2000; 2000US-189320P.
PR      15-MAR-2000; 2000US-189328P.
PR      21-MAR-2000; 2000US-190828P.
PR      21-MAR-2000; 2000US-191007P.
PR      21-MAR-2000; 2000US-191048P.
PR      21-MAR-2000; 2000US-191314P.
PR      28-MAR-2000; 2000US-192655P.
PR      29-MAR-2000; 2000US-193032P.
PR      30-MAR-2000; 2000US-193053P.
PR      04-APR-2000; 2000WO-US08439.
PR      04-APR-2000; 2000US-194449P.
PR      11-APR-2000; 2000US-194647P.
PR      11-APR-2000; 2000US-195975P.
PR      11-APR-2000; 2000US-196000P.
PR      11-APR-2000; 2000US-196690P.
PR      11-APR-2000; 2000US-196820P.
PR      18-APR-2000; 2000US-198121P.
PR      18-APR-2000; 2000US-198585P.
PR      25-APR-2000; 2000US-199397P.
PR      25-APR-2000; 2000US-199550P.
PR      25-APR-2000; 2000US-199654P.
PR      03-MAY-2000; 2000US-201516P.
PR      17-MAY-2000; 2000WO-US13705.
PR      22-MAY-2000; 2000WO-US14042.
PR      30-MAY-2000; 2000WO-US14941.
PR      02-JUN-2000; 2000WO-US15264.
PR      05-JUN-2000; 2000US-209832P.
PR      28-JUL-2000; 2000WO-US20710.
PR      22-AUG-2000; 2000US-0644848.
PR      24-AUG-2000; 2000WO-US23328.
PR      08-NOV-2000; 2000WO-US30952.
PR      01-DEC-2000; 2000WO-US32678.
PR      20-DEC-2000; 2000WO-US34956.
XX
PA      (GETH ) GENENTECH INC.
XX
PI      Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI      Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX
XX      WPI: 2001-602746/68.
XX      P-PSDB; AAU29227.
XX
PT      Novel nucleic acids encoding PRO polypeptides, used to diagnose the
PT      presence of tumours, such as prostate and breast tumours, in mammals and

```

```

QY 781 TGGCAAGACATCAAGATGTGACACAGAACACAGTGCACCTGCAGACAGAGCTGC 840
Db 1237 TGGCAAGACATCAAGATGTGACACAGAACACAGTGCACCTGCAGACAGAGCTGC 1296
QY 841 TACAATTTACAAAGGGGCTTCAAAATGCATGCAGACCCCATCCGCTGTGAGAGGCTTATCTG 900
Db 1297 TACAATTTACAAAGGGGCTTCAAAATGCATGCAGACCCCATCCGCTGTGAGAGGCTTATCTG 1356
QY 901 AGGATCATGTATTAACCCCTGTATGTCTCTGTGAGAACCCCTGGCTGCAGAACACAGCC 960
Db 1357 AGGATCATGTATTAACCCCTGTATGTCTCTGTGAGAACCCCTGGCTGCAGAACACAGCC 1416
QY 961 TTTACCATCTTGTACCGGACATGAGACGTGTGTGACAGACGCTCCGTTCCGCTGCATC 1020
Db 1417 TTTACCATCTTGTACCGGACATGAGACGTGTGTGACAGACGCTCCGTTCCGCTGCATC 1476
QY 1021 TTTCAATTCAGACGACGACCCGCTACCCCTGGGCTTATACATTTTCCAGATCAATCT 1080
Db 1477 TTTCAATTCAGACGACGACCCGCTACCCCTGGGCTTATACATTTTCCAGATCAATCT 1536
QY 1081 GGGAAATGAGGAGAGATTTTACATGGGGGCAAAAGGGGCCCCATGATGTCACCCCTGCTG 1140
Db 1537 GGGAAATGAGGAGAGATTTTACATGGGGGCAAAAGGGGCCCCATGATGTCACCCCTGCTG 1596
QY 1141 ATGACAGCGCCCATCAAAAGGGGCCCCGGAATCCAGCTGACCTTGAATGATCACTGTC 1200
Db 1597 ATGACAGCGCCCATCAAAAGGGGCCCCGGAATCCAGCTGACCTTGAATGATCACTGTC 1656
QY 1201 AACACTGTCAATCACTTCAGAGGACGCTCGTGTATCCGATCGCATATATGTGTGCGAG 1260
Db 1657 AACACTGTCAATCACTTCAGAGGACGCTCGTGTATCCGATCGCATATATGTGTGCGAG 1716
QY 1261 TACCATTTC 1269
Db 1717 TACCATTTC 1725

```

RESULT 9
AA37670 standard; cDNA; 2509 BP.

AA37670:
11-SEP-2000 (first entry)

Human EGF-like homologue (PRO217) encoded by DNA32279 cDNA.

KW Inflammatory cell infiltration; immune response; T cell proliferation;
 KW anti-inflammatory; anti-autoimmune; anti-diabetic; spondyloarthropathy;
 KW T cell-mediated disease; spondyloarthropathy; sclerosis; renal disease;
 KW inflammatory myopathy; hemolytic anemia; thrombocytopenia; thyroiditis;
 KW diabetes mellitus; demyelinating polyneuropathy; Guillain-Barre syndrome;
 KW multiple sclerosis; polyneuropathy; hepatitis; cirrhosis; enteropathy;
 KW sclerosing cholangitis; inflammatory bowel disease; Whipple's disease;
 KW skin disease; dermatitis; psoriasis; asthma; allergic rhinitis; tumor;
 KW food hypersensitivity; urticaria; eosinophilic pneumonia; transplant;
 KW idiopathic pulmonary fibrosis; graft rejection; PRO245; human; PRO217;
 KW EGF-like; ss.

OS Homo sapiens.
 XX
 XX
 XX MO9914241-A2.
 PD
 PD 25-MAR-1999.
 PE 17-SEP-1998; 98MO-US19437.
 XX
 XX 17-SEP-1997; 97US-0059119.
 PR 17-SEP-1997; 97US-0059119.
 PR 18-SEP-1997; 97US-0059263.
 PR 28-OCT-1997; 97US-0063550.
 PR 12-NOV-1997; 97US-0065186.
 PR 21-NOV-1997; 97US-0066364.

```

PR 24-NOV-1997; 97US-0066770.  

PR 04-JUN-1998; 98US-0088026.  

XX  

PA (GETH ) GENENTECH INC.  

XX  

PI Fong S, Goddard A, Gurney AL, Tumas D, Wood WI;  

XX  

DR WPI, 1999-229499/19.  

XX  

PT Composition containing novel polypeptide PRO245, its agonist or  

PT antagonist -  

PS  

PS Example 1: Fig 5A-1-5A-2; 177pp; English.  

XX  

CC This invention describes a novel composition containing (apart from a  

CC carrier or excipient), a novel PRO245 polypeptide (I), its agonist or  

CC antagonist, or their fragments, for modulating: (i) infiltration of  

CC inflammatory cells into tissue; (ii) an immune response; or (iii) T cell  

CC proliferation. The composition increases or decreases any of the effects  

CC (i)-(iii). The products of the invention have anti-inflammatory,  

CC anti-autoimmune and anti-diabetic activity. (I), and its (ant)agonists  

CC and their fragments, are used to treat immune-related diseases,  

CC particularly T cell-mediated diseases. The diseases treated include  

CC systemic lupus erythematosus, rheumatoid arthritis, juvenile chronic  

CC arthritis, spondyloarthropathies, systemic sclerosis (scleroderma),  

CC idiopathic inflammatory myopathies (dermatomyositis, polymyositis),  

CC Sjogren's syndrome, systemic vasculitis, sarcoidosis, autoimmune  

CC hemolytic anemia (immune pancytopenia, paroxysmal nocturnal  

CC hemoglobinuria), autoimmune thrombocytopenia (idiopathic thrombocytopenic  

CC purpura immune-mediated thrombocytopenia), thyroiditis (Grave's disease,  

CC Hashimoto's thyroiditis, juvenile lymphocytic thyroiditis, atrophic  

CC thyroiditis), diabetes mellitus, immune-mediated renal disease  

CC (glomerulonephritis, tubulointerstitial nephritis), multiple sclerosis,  

CC idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome, chronic  

CC inflammatory demyelinating polyneuropathy, infectious hepatitis  

CC (hepatitis A, B, C, D, E and other non-hepatotropic viruses), autoimmune  

CC chronic active hepatitis, primary biliary cirrhosis, granulomatous  

CC hepatitis, and sclerosing cholangitis, inflammatory bowel disease  

CC (ulcerative colitis; Crohn's disease), gluten-sensitive enteropathy, and  

CC Whipple's disease. Autoimmune or immune-mediated skin diseases including  

CC bullous skin diseases, erythema multiforme, contact dermatitis, psoriasis,  

CC asthma, allergic rhinitis, atopic dermatitis, food hypersensitivity,  

CC urticaria, eosinophilic pneumonia, idiopathic pulmonary fibrosis,  

CC hypersensitivity pneumonitis, and transplant rejection associated diseases  

CC (graft rejection, and graft-versus-host-disease). (I), its (ant)agonists  

CC or fragment can also be used as an adjuvant in treatment of tumors.  

CC Antibodies against (I) can also be used for diagnosing such diseases.  

CC This sequence encodes a human EGF-like homologue (PRO217) encoded by cDNA  

CC clone DNA32279 which is described in the invention.  

XX  

SQ Sequence 2509 BP; 591 A; 695 C; 590 G; 633 T; 0 other;

```

Query Match 99.9%; Score 1267.4; DB 20; Length 2509;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 CAGTGCAGATGAGCTTGTACCTGGATGCGCAGTGCAGACAGTGTATGATATGATGA 60
Db 525 CAGTGCAGATGAGCTTGTACCTGGATGCGCAGTGCAGACAGTGTATGATATGATGA 584
QY 61 TGGCGAACATCCCGAGGCGCTGCGGAGACAGATGATGTGTAAACCAAAATGCGGG 120
Db 585 TGGCGAACATCCCGAGGCGCTGCGGAGACAGATGATGTGTAAACCAAAATGCGGG 644
QY 121 TATTATGATTCCTCCCGGACAAACCTGTGTATGAGGCGCTTACCGAAACCCCTACTCG 180
Db 645 TATTATGATTCCTCCCGGACAAACCTGTGTATGAGGCGCTTACCGAAACCCCTACTCG 704
QY 181 ACCCCCTACTGAGTCCCTTACCCAGACAGCTGCCCCACAGCTCTCAAGTATGCC 240
Db 705 ACCCCCTACTGAGTCCCTTACCCAGACAGCTGCCCCACAGCTCTCAAGTATGCC 764
QY 241 ACGATCTCCAGGCTCTTATATGCCGCTTGTGATACAGATGATGAAGCAACCAATGT 300

```

```

OY 721 CCCGACATATCTTCTGCTCTGCCCTCCAGGCTACATCTCTGTGATGACACCGAAC 780
DB 981 CCCGGACATATCTTCTGCTCTGCCCTCCAGGCTACATCTCTGTGATGACACCGAAC 1040
OY 781 TGGCAGACATCAACGATGTGAGACAGAACACACAGTGCACCTGAGAGAGAGTGC 840
DB 1041 TGGCAGACATCAACGATGTGAGACAGAACACACAGTGCACCTGAGAGAGAGTGC 1100
OY 841 TACAAATTACAGGGGCTTCAATGCAATGCAACCCATCCGCTGTGAGAGAGCTTATCTG 900
DB 1101 TACAAATTACAGGGGCTTCAATGCAATGCAACCCATCCGCTGTGAGAGAGCTTATCTG 1160
OY 901 AGGATCAGTATAACCGCTGTATGTGCTGCTGAGAACCCGCTGAGAGAGAGAGAGCC 960
DB 1161 AGGATCAGTATAACCGCTGTATGTGCTGCTGAGAACCCGCTGAGAGAGAGAGAGCC 1220
OY 961 TTTACCATCTTGTACCGGGACATGAGAGTGTGTCAGAGACGCTCCGCTGACATC 1020
DB 1221 TTTACCATCTTGTACCGGGACATGAGAGTGTGTCAGAGACGCTCCGCTGACATC 1280
OY 1021 TTCCAATGCAAGCGACGACCCGCTACCCCTGAGGCTTATTCATTTTCCAGATCAAACT 1080
DB 1281 TTCCAATGCAAGCGACGACCCGCTACCCCTGAGGCTTATTCATTTTCCAGATCAAACT 1340
OY 1081 GGGATGAGGCGAGAGATTTTACATGCGGCAACGGGCCCCATCAGTGCACCCCTGCTG 1140
DB 1341 GGGATGAGGCGAGAGATTTTACATGCGGCAACGGGCCCCATCAGTGCACCCCTGCTG 1400
OY 1141 ATGACACGCGCCCATCAAGGGCCCCGGGAAATCCAGCTGCACTTGGAATGATCAGTCTC 1200
DB 1401 ATGACACGCGCCCATCAAGGGCCCCGGGAAATCCAGCTGCACTTGGAATGATCAGTCTC 1460
OY 1201 AACACTGTATCAACTTCAAGAGAGCTCCGTGATCGAGCGGGATATATGTGCTGAG 1260
DB 1461 AACACTGTATCAACTTCAAGAGAGCTCCGTGATCGAGCGGGATATATGTGCTGAG 1520
OY 1261 TACCCATTC 1269
DB 1521 TACCCATTC 1529

```

RESULT 8
AAK05359
ID AAK05359 standard; DNA; 2550 BP.
XX AAK05359;
AC
XX
XX 07-MAY-1999 (first entry)
DE Human extracellular matrix protein (ECMP)-1 encoding DNA.
XX
XX Extracellular matrix protein; ECMP-1; diagnosis; treatment; cancer;
XX Immune disorder; human; ss.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX CDS 382..1728
XX FT /tag="a
XX FT /product="ECMP-1"
XX
XX MO9900410-A2.
XX
XX 07-JAN-1999.
XX
XX 23-JUN-1998; 98MO-US13012.
XX
XX 27-JUN-1997; 97US-0884072.
XX (INCY-) INCYTE PHARM INC.
XX
XX Bandman O, Corley NC, Guegler KJ;
PI

```

XX WI: 1999-095674/08.
XX P-PSDB; AAM94281.
XX
XX New polynucleotide encoding extracellular matrix protein, ECMP-1 -
XX useful in the diagnosis, prevention and treatment of immune
XX disorders and cancer
XX
XX Claim 3: Fig 1A-G; 79pp; English.
XX
XX This DNA encodes a human extracellular matrix protein (ECMP)-1. Host
XX cells containing a vector comprising the ECMP-1 nucleic acid are used
XX for the recombinant production of the protein. ECMP-1 and its
XX (ant)agonists, are useful in the diagnosis, prevention, and treatment
XX of cancer and immune disorders.
XX
XX Sequence 2550 BP; 623 A; 690 C; 596 G; 641 T; 0 other:
XX
XX Query Match 100.0%; Score 1269; DB 20; Length 2550;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 CAGTGCAGCATGAGCTTGTGACCTGATGCGGACGACGACGATTTAGATTTAGTAA 60
DB 457 CAGTGCAGCATGAGCTTGTGACCTGATGCGGACGACGACGATTTAGATTTAGTAA 516
OY 61 TGGCAGAACATCCCGAGGCTGCGGAGAGACATGATGTGTAAACAAATGGCGGG 120
DB 517 TGGCAGAACATCCCGAGGCTGCGGAGAGACATGATGTGTAAACAAATGGCGGG 576
OY 121 TATTATGATTCCTCCCGGCAAAACCTGTGTATCGAGGCGCTTACGAAACCCCTACTCG 180
DB 577 TATTATGATTCCTCCCGGCAAAACCTGTGTATCGAGGCGCTTACGAAACCCCTACTCG 636
OY 181 ACCCCCTACTCAGTCCGATCCAGACAGAGCTGCCACCACTCTCAGCTCCAAACTATCC 240
DB 637 ACCCCCTACTCAGTCCGATCCAGACAGAGCTGCCACCACTCTCAGCTCCAAACTATCC 696
OY 241 ACGATCTCCAGGCTCTTATATATGCGGCTTGGATACAGATGAGTAAACCAATGT 300
DB 697 ACGATCTCCAGGCTCTTATATATGCGGCTTGGATACAGATGAGTAAACCAATGT 756
OY 301 GTGATGTGACGAGTGTGCAACAGATTCACACAGTGAACCCCAACCAAGATCTGATC 360
DB 757 GTGATGTGACGAGTGTGCAACAGATTCACACAGTGAACCCCAACCAAGATCTGATC 816
OY 361 AATACGTAAGGCGGCTTACACCTGCTGACACGAGATATTTGGCTTCTGGAAGCCAG 420
DB 817 AATACGTAAGGCGGCTTACACCTGCTGACACGAGATATTTGGCTTCTGGAAGCCAG 876
OY 421 TGCCTAGACATGATGATGCGTATGCTTACTGTCGACAGACTCTGTCCGAATGTCT 480
DB 877 TGCCTAGACATGATGATGCGTATGCTTACTGTCGACAGACTCTGTCCGAATGTCT 936
OY 481 GGATCCATTTCTTGTACATGCAACCCGTTTACCTCAATGAGATGGAAGTCTTATGC 540
DB 937 GGATCCATTTCTTGTACATGCAACCCGTTTACCTCAATGAGATGGAAGTCTTATGC 996
OY 541 CAAGATGTGAACGAGTGTGCCACGAGAACCCCTGCTGCAAACTCGCTCAACACTAC 600
DB 997 CAAGATGTGAACGAGTGTGCCACGAGAACCCCTGCTGCAAACTCGCTCAACACTAC 1056
OY 601 GGCTCTTTCATCTGCGGCTGTGACCCAGATATGAACCTTGGAAGATGGGCTTATGC 660
DB 1057 GGCTCTTTCATCTGCGGCTGTGACCCAGATATGAACCTTGGAAGATGGGCTTATGC 1116
OY 661 AGTGAATGACAGAGTGCAGCTTCTGAGTTCCTGCAACATGAGTGTGTAACAG 720
DB 1117 AGTGAATGACAGAGTGCAGCTTCTGAGTTCCTGCAACATGAGTGTGTAACAG 1176
OY 721 CCCGGACATATCTTCTGCTCTGCCCTCCAGGCTACATCTCTGTGATGACACCGAAC 780
DB 1177 CCCGGACATATCTTCTGCTCTGCCCTCCAGGCTACATCTCTGTGATGACACCGAAC 1236

```

Db 1281 TTCCAAATGCAAGCCAGCAGCCGCTACCTGGGGCTATTACATTTTCCAGATCAAAATCT 1340
 Qy 1081 GGGATGAGGGGAGAGATTTTACATGGGCAAGGGCCCATAGTGGCCCTGGTG 1140
 Db 1341 GGGATGAGGGGAGAGATTTTACATGGGCAAGGGCCCATAGTGGCCCTGGTG 1400
 Qy 1141 ATGACAGCGCCCATCAAAAGGGCCCGGAAATCCAGCTGTGAAATGATCACTGTC 1200
 Db 1401 ATGACAGCGCCCATCAAAAGGGCCCGGAAATCCAGCTGTGAAATGATCACTGTC 1460
 Qy 1201 AACACTGTCACTCACTTCAAGAGCAGCTCCGTATCCGACGTGGGATATGTGTGCGAG 1260
 Db 1461 AACACTGTCACTCACTTCAAGAGCAGCTCCGTATCCGACGTGGGATATGTGTGCGAG 1520
 Qy 1261 TACCATTC 1269
 Db 1521 TACCATTC 1529

RESULT 7

AA39892
 AA39892 standard; cDNA; 2362 BP.

AA39892:

21-FEB-2000 (first entry)

cDNA sequence of the human secreted protein AK647.

AK647: aortic tissue development; smooth muscle cell modulator; SCID;
 nutritional supplement; vasculogenesis; embryonic development; infection;
 cytokine activity; cell proliferation; cell differentiation; defect; HIV;
 immune deficiency; haematopoiesis regulation; tissue regrowth; diagnosis;
 wound healing; restenosis; atherosclerosis; drug screen; ss.

Homo sapiens.

Key Location/Qualifiers
 CDS 185..1532
 FT /*lag- a
 FT /product- AK647
 FT /note- "Secreted protein"

W09960125-A2.

25-NOV-1999.

18-MAY-1999; 99MO-US10931.

19-MAY-1998; 98US-0081002.

21-MAY-1998; 98US-0083002.

(GEMV) GENETICS INST INC.

Jacobs K, McCoy JM, Racie L, Lavallie E, Treacy M, Evans C;
 Agostino M, Lu Z, Merberg D;

WPI; 2000-053298/04.

P-PSDB; AAY57058.

Proteins, and their encoding polynucleotides, used for treating e.g.
 smooth muscle cell growth, vasculogenesis, restenosis or
 atherosclerosis

Claim 1; Page 45-46; 49pp; English.

This is the nucleotide sequence of the human secreted protein AK647. The
 polynucleotide sequence was obtained from a human foetal kidney cDNA
 library. AK647 homologues in chicks and rodents are involved in aortic
 tissue development. The spatial and temporal distribution of AK647
 indicated that it acts as a modulator of smooth muscle cells in
 vasculogenesis during embryonic development. The primary structure of
 AK647 consists of multiple EGF domains. The AK647 protein can be used as

CC a nutritional source or supplement. The protein shows both inhibitory and
 CC inducing, cytokine, cell proliferation and cell differentiation activity.
 CC The protein may also be used in the treatment of immune deficiencies and
 CC disorders, including severe combined immunodeficiency (SCID), HIV and
 CC other viral, bacterial and fungal infections. Regulation of immune
 CC responses may also be carried out by the AK647 protein. Other uses of the
 CC protein include a role in the regulation of haematopoiesis and in the
 CC treatment of myeloid and lymphoid cell deficiencies. Uses in bone,
 CC cartilage, tendon, ligament and nerve tissue regrowth are also possible,
 CC as well as for wound healing and in the treatment of ulcers and burns.
 CC The polynucleotides and proteins can be used for preventing, treating or
 CC ameliorating smooth muscle cell growth, vasculogenesis, restenosis,
 CC atherosclerosis, blood vessel remodelling and degeneration. The proteins
 CC may also have activin/inhibin, chemotactic/chemokinetic, haemostatic and
 CC thrombolytic, receptor/ligand, anti-inflammatory, cadherin/tumour
 CC invasion suppressor, and tumour inhibition activity. AK647 specific
 CC antibodies can be used for promoting smooth muscle cell growth or
 CC vasculogenesis. The proteins and polynucleotides can also be used for
 CC detection, diagnosis and drug screening.

Sequence 2362 BP; 580 A; 650 C; 548 G; 584 T; 0 other;

Query Match 100.0%; Score 1269; DB 21; Length 2362;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 1269; Conservative 0; Indels 0; Gaps 0;

Qy 1 CAGTGCAGATAGGCTTGGACCTGGATGCGCCAGTGCAGAGATGTTAGATATGATGA 60
 Db 261 CAGTGCAGATAGGCTTGGACCTGGATGCGCCAGTGCAGAGATGTTAGATATGATGA 320
 Qy 61 TGGCGAACCATCCCGAGGCGCTGGCGAGAGACATGATGTGTTAACCAAAATGGCGGG 120
 Db 321 TGGCGAACCATCCCGAGGCGCTGGCGAGAGACATGATGTGTTAACCAAAATGGCGGG 380
 Qy 121 TATTATGACATTCGCCGAGCAAAACCTGTGATGAGAGGCGCTACTGCAACCCCTACTG 180
 Db 381 TATTATGACATTCGCCGAGCAAAACCTGTGATGAGAGGCGCTACTGCAACCCCTACTG 440
 Qy 181 ACCCGTACATCAGTCCGTACCCAGCAGCTGCCGCCACCACTCAGCTCAAACTATCC 240
 Db 441 ACCCGTACATCAGTCCGTACCCAGCAGCTGCCGCCACCACTCAGCTCAAACTATCC 500
 Qy 241 ACGATCTCAGGCGCTTTATATGCGCGTTGGATACCAAGATGATGAACCAACCAATGT 300
 Db 501 ACGATCTCAGGCGCTTTATATGCGCGTTGGATACCAAGATGATGAACCAACCAATGT 560
 Qy 301 GTGATGTGACAGAGTGTGCAACAGATTCGCCACAGTGCACACCCACCAACCAATGTGCATC 360
 Db 561 GTGATGTGACAGAGTGTGCAACAGATTCGCCACAGTGCACACCCACCAACCAATGTGCATC 620
 Qy 361 AATACGAGAGGCGGTACACCTGCTGCTGCACCGAGGATATGTGGCTTGTGAAAGGCCAG 420
 Db 621 AATACGAGAGGCGGTACACCTGCTGCTGCACCGAGGATATGTGGCTTGTGAAAGGCCAG 680
 Qy 421 TGGTTAGACATTTGATGATGTGCTATGCTTACTCCAGCAGCTGTGCGAATGTCTCT 480
 Db 681 TGGTTAGACATTTGATGATGTGCTATGCTTACTCCAGCAGCTGTGCGAATGTCTCT 740
 Qy 481 GGATCCATATCTTATACATGCAACCCGTGTTTACCTCAATGAGAGTGAAGGTCTTGC 540
 Db 741 GGATCCATATCTTATACATGCAACCCGTGTTTACCTCAATGAGAGTGAAGGTCTTGC 800
 Qy 541 CAAGATGTGAACGAGTGTGCCACGAGAACCCCTGCTGCAAAACCTGCTCAACACCTTAC 600
 Db 801 CAAGATGTGAACGAGTGTGCCACGAGAACCCCTGCTGCAAAACCTGCTCAACACCTTAC 860
 Qy 601 GGCTCTTCAATCTCCGCTGTGACCCAGATATGAACTTGAGGAAGATGGCTTCATTCG 660
 Db 861 GGCTCTTCAATCTCCGCTGTGACCCAGATATGAACTTGAGGAAGATGGCTTCATTCG 920
 Qy 661 AGTATATGACAGAGTGCAGCTTCTGTAGTTCCTCTGCAACATGAGTGTGAACGAG 720
 Db 921 AGTATATGACAGAGTGCAGCTTCTGTAGTTCCTCTGCAACATGAGTGTGAACGAG 980

Db 1324 GGGATGAGGCGAGAGATTTTACATGCGGCAAAACGGGCCCCCATCATGTCACACCCGCGTG 1383
 QY 1141 ATGACAGCCCCCATCAAAAGGCCCCGGGAATCCAGCTGGAGATTTGGAATGATCACTGTC 1200
 Db 1384 ATGACAGCCCCCATCAAAAGGCCCCGGGAATCCAGCTGGAGATTTGGAATGATCACTGTC 1443
 QY 1201 AACACTGTCATCAACTTCAGAGGAGCTCCGATCCAGCTGGAGATTTGATGTCGCGAG 1260
 Db 1444 AACACTGTCATCAACTTCAGAGGAGCTCCGATCCAGCTGGAGATTTGATGTCGCGAG 1503
 QY 1261 TACCCATTC 1269
 Db 1504 TACCCATTC 1512

RESULT 6
 AAX07567
 ID AAX07567 standard; cDNA: 2362 BP.
 AAX07567:
 21-JUN-1999 (first entry)
 Homo sapiens fetal kidney clone AK647 secreted protein gene.
 Secreted protein: fetal kidney; ds.
 Homo sapiens.
 Key location/Qualifiers
 CDS 186..1532
 FT /tag= a
 FT /product= "secreted protein"
 W09900405-A1.
 07-JAN-1999.
 29-JUN-1998: 98MO-US13530.
 30-JUN-1997: 97US-0885610.
 (GENY) GENETICS INST INC.
 Agostino M, Evans C, Jacobs K, Lavallie ER, McCoy JM;
 Merderg D, Racie LA, Treacy M;
 MPI: 1999-095671/08.
 P-PSDB: AAW95709.
 New polynucleotides encoding secreted human proteins - are derived
 from fetal kidney or adult retina cDNA libraries, used as, e.g.
 potential vaccines
 Disclosure: Pages 51-52; 76pp; English.

The sequence is that encoding a secreted protein from a human fetal
 kidney clone AK647. Such a sequence is predicted to have biological
 activities which would make them suitable for treating, preventing or
 ameliorating medical conditions in humans and animals, although no
 supporting data is given. Suggested activities include nutritional
 activity, cytokine and cell proliferation/differentiation activity,
 immune stimulating (e.g. as vaccines) or suppressing activity,
 haematopoiesis regulating activity, tissue growth activity,
 activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 and thrombolytic activity, receptor/ligand activity, anti-inflammatory
 activity, cadherin/tumour invasion suppressor activity, and tumour
 inhibition activity. It is also stated to be useful for gene
 therapy.

Sequence 2362 BP; 580 A; 650 C; 548 G; 584 T; 0 other;

Query Match 100.0%; Score 1269; DB 20; Length 2362;

Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGTGCAGCAATGGCTTGTACCTGATCGGACGACGACAGTGTATGATGAA 60
 Db 261 CAGTGCAGCAATGGCTTGTACCTGATCGGACGACGACAGTGTATGATGAA 320
 QY 61 TGGCGAACCATCCCGAGGCGCTCCGAGAGACATGATGTGTAAACAAATGGCGG 120
 Db 321 TGGCGAACCATCCCGAGGCGCTCCGAGAGACATGATGTGTAAACAAATGGCGG 380
 QY 121 TATTTATGATCCCGGACAAACCTGTGTATCGAGGCGCTTACTGAAACCTTCTCG 180
 Db 381 TATTTATGATCCCGGACAAACCTGTGTATCGAGGCGCTTACTGAAACCTTCTCG 440
 QY 181 ACCCCCTACTCAGTCCGATACCGAGAGCTGCCACACTCTCAGCTCCAACTATGCC 240
 Db 441 ACCCCCTACTCAGTCCGATACCGAGAGCTGCCACACTCTCAGCTCCAACTATGCC 500
 QY 241 ACGATCTCCAGGCTCTTATATGCGCGCTTGTATACCATGATGAAACCAATGT 300
 Db 501 ACGATCTCCAGGCTCTTATATGCGCGCTTGTATACCATGATGAAACCAATGT 560
 QY 301 GTGATGTGAGAGATGTGCAACAGATTCGCCACAGTGCACACCCCAACCCAGATCTGATC 360
 Db 561 GTGATGTGAGAGATGTGCAACAGATTCGCCACAGTGCACACCCCAACCCAGATCTGATC 620
 QY 361 AATACTGAAGGGGGGTACACCTGCTCCACCGACGATATGGCTTGTGAAGGCGAG 420
 Db 621 AATACTGAAGGGGGGTACACCTGCTCCACCGACGATATGGCTTGTGAAGGCGAG 680
 QY 421 TGCTTAGACATGATGAATGGTGGCTATGCTTACCTGCACAGCTCTGTGGAATGTTCCT 480
 Db 681 TGCTTAGACATGATGAATGGTGGCTATGCTTACCTGCACAGCTCTGTGGAATGTTCCT 740
 QY 481 GGATCCTATTTCTGTACATGCAACCTGGTTTACCTCAATGAGATGGAAGGCTTTCG 540
 Db 741 GGATCCTATTTCTGTACATGCAACCTGGTTTACCTCAATGAGATGGAAGGCTTTCG 800
 QY 541 CAAAGTGTGAACGAGTGTGCCACGAGAAACCCCTGGTGCAAACTGCTCAACACTTAC 600
 Db 801 CAAAGTGTGAACGAGTGTGCCACGAGAAACCCCTGGTGCAAACTGCTCAACACTTAC 860
 QY 601 GGCTCTTCACTGCGCTGTGACCCAGATATGAATCTGAGAGATGGCTTCAATTGC 660
 Db 861 GGCTCTTCACTGCGCTGTGACCCAGATATGAATCTGAGAGATGGCTTCAATTGC 920
 QY 661 AGTATATGACGAGTGTGCTTCTGTAGTTCCTGTGCCAATGATGATGTGAACGAG 720
 Db 921 AGTATATGACGAGTGTGCTTCTGTAGTTCCTGTGCCAATGATGATGTGAACGAG 980
 QY 721 CCGCGCAATCTTCTGCTGCTGCTCCAGGCTCAATCTGCTGTGATGACCAACGAGC 780
 Db 981 CCGCGCAATCTTCTGCTGCTGCTCCAGGCTCAATCTGCTGTGATGACCAACGAGC 1040
 QY 781 TGCAGACATCAAGATGTGAGACAGAGAACCAAGTGAACCTGTGAGAGACAGCGTC 840
 Db 1041 TGCAGACATCAAGATGTGAGACAGAGAACCAAGTGAACCTGTGAGAGACAGCGTC 1100
 QY 841 TACATTTTACAGGGGGCTTCAAAATGATGACACCCCATCCGTGTGAGAGAGCTTATCTG 900
 Db 1101 TACATTTTACAGGGGGCTTCAAAATGATGACACCCCATCCGTGTGAGAGAGCTTATCTG 1160
 QY 901 AGGATCAGTGTAAACGCTGTATGTGTCCTGTAGAACCTGTGTGAGAGACAGCGCC 960
 Db 1161 AGGATCAGTGTAAACGCTGTATGTGTCCTGTAGAACCTGTGTGAGAGACAGCGCC 1220
 QY 961 TTACCAATCTTGTACCGGAGCATGAGAGTGTGTACAGAGACCTCGTCCGCTGACATC 1020
 Db 1221 TTACCAATCTTGTACCGGAGCATGAGAGTGTGTACAGAGACCTCGTCCGCTGACATC 1280
 QY 1021 TTCCAAATGCAAGCAGACCCCGCTACCTGGGGCTTATTAATTTTCAGATCAAAATCT 1080

[illegible]

D	B	244	CAGTGCAGGATGGCTTTGACCTGGATGCCAGTCCAGCAGTGTGTTTAAACCAAAATGGCGGG	120
O	Y	61	TGCGGAACCATCCCGAGGCGCTGCCGAGAGACATGATGTGTAAACCAAAATGGCGGG	120
D	B	304	TGCCGAACCATCCCGAGGCGCTGCCGAGAGACATGATGTGTAAACCAAAATGGCGGG	363
O	Y	121	TATTTATGCAATTCGCCGGAACAAACCCCTGTATGTGAGGGGCCCTACTCGAACCCCTACTGC	180
D	B	364	TATTTATGCAATTCGCCGGAACAAACCCCTGTATGTGAGGGGCCCTACTCGAACCCCTACTGC	423
O	Y	181	ACCCCTACTGAGGCGGTACCCAGCAGCTCCCGACCATCTCAGCTCCAGCTCAAACTATCC	240
D	B	424	ACCCCTACTGAGGCGGTACCCAGCAGCTCCCGACCATCTCAGCTCCAGCTCAAACTATCC	483
O	Y	241	ACGATCTCCAGGCGCTTTATATGCGCGCTTTGGATACCAATGATGATAAAGCAACCAATGT	300
D	B	484	ACGATCTCCAGGCGCTTTATATGCGCGCTTTGGATACCAATGATGATAAAGCAACCAATGT	543
O	Y	301	GTGATGTGAGAGAGTGTGCAACAGATTCCACACGTCACACCCAGCCAGATCTGCATTC	360
D	B	544	GTGATGTGAGAGAGTGTGCAACAGATTCCACACGTCACACCCAGCCAGATCTGCATTC	603
O	Y	361	AATACGTAAGGCGGGTACACCTGCTCTGCAACCGACGAGATATGTGGCTTCTGGAAGCCAG	420
D	B	604	AATACGTAAGGCGGGTACACCTGCTCTGCAACCGACGAGATATGTGGCTTCTGGAAGCCAG	663
O	Y	421	TGCTTAGCATTTGATGATGTGCGTATGCGTACTGCGACGAGCTGTGCGCAATGTCTCT	480
D	B	664	TGCTTAGCATTTGATGATGTGCGTATGCGTACTGCGACGAGCTGTGCGCAATGTCTCT	723
O	Y	481	GGATTCATTCCTGTACATGCAACCCCTGGTTTAACTCAATGAGAGTGAAGCTCTTGC	540
D	B	724	GGATTCATTCCTGTACATGCAACCCCTGGTTTAACTCAATGAGAGTGAAGCTCTTGC	783
O	Y	541	CAAGATGGAAGAGTGTGCCACCGAGAAACCCCTGCGTCAAACTGGGTCAACACTAC	600
D	B	784	CAAGATGGAAGAGTGTGCCACCGAGAAACCCCTGCGTCAAACTGGGTCAACACTAC	843
O	Y	601	GGCTCTTTCATCTCGCGCTGTGACCCAGCAAGATATGAACTTGAGGAAGATGGGCTCATTC	660
D	B	844	GGCTCTTTCATCTCGCGCTGTGACCCAGCAAGATATGAACTTGAGGAAGATGGGCTCATTC	903
O	Y	661	AGTATATGAGACAGATGACGCTTCTCTGAGTTCTCTGCGCAACATGATGTGTGACAG	720
D	B	904	AGTATATGAGACAGATGACGCTTCTCTGAGTTCTCTGCGCAACATGATGTGTGACAG	963
O	Y	721	CCCGGCAATACTTCTGCTCTGCGCCCTCAGGCTCAATCTGCTGAGTGAACAACCGAAGC	780
D	B	964	CCCGGCAATACTTCTGCTCTGCGCCCTCAGGCTCAATCTGCTGAGTGAACAACCGAAGC	1022
O	Y	781	TGCCAAGACATCAAGAAATGTGAGCACAGAAACCAACAGTGCMAACTTCACAGACAGCTGC	840
D	B	1024	TGCCAAGACATCAAGAAATGTGAGCACAGAAACCAACAGTGCMAACTTCACAGACAGCTGC	1083
O	Y	841	TACATTTTACAAAGGGGCTTCAAAATGCAATGCAACCCATCCGCTGTGAGAGAGCTTATGTG	900
D	B	1084	TACATTTTACAAAGGGGCTTCAAAATGCAATGCAACCCATCCGCTGTGAGAGAGCTTATGTG	1143
O	Y	901	AGGATCAGATGAACCCGCTGTATGTGCTCTGAGAACCCCTGGCTGACAGAGACCGCC	960
D	B	1144	AGGATCAGATGAACCCGCTGTATGTGCTCTGAGAACCCCTGGCTGACAGAGACCGCC	1203
O	Y	961	TTTACCATCTTGTACCGGAGACATGAGCGTGTGACGACGCTCGTTCGCGCTGACATC	1020
D	B	1204	TTTACCATCTTGTACCGGAGACATGAGCGTGTGACGACGCTCGTTCGCGCTGACATC	1263
O	Y	1021	TTTCAAAATGCAAGCACGACCGCGTACCCCTGGGGGCTTTACATTTTCCAGATCAAAATCT	1080
D	B	1264	TTTCAAAATGCAAGCACGACCGCGTACCCCTGGGGGCTTTACATTTTCCAGATCAAAATCT	1323
O	Y	1081	GGGATGAGGGCAGAGAAATTTTACATGCGGCAACCGGGCCCATCAGTGCACCCCTGGTG	1140

QY 1141 ATGACAGCCCCCAAGGGCCCCGGAATTCAGCTGAGTGGAAATGATCACTGTC 1200
DB 1426 ATGACAGCCCCCAAGGGCCCCGGAATTCAGCTGAGTGGAAATGATCACTGTC 1485
QY 1201 AACACTGTCACTCACTCAGAGGAGCTCCGTGATCCGACTGCGATATATGTGTCGAG 1260
DB 1486 AACACTGTCACTCACTCAGAGGAGCTCCGTGATCCGACTGCGATATATGTGTCGAG 1545
QY 1261 TACCATTC 1269
DB 1546 TACCATTC 1554

RESULT 4
AAK94505
ID AAK94505 standard; cDNA: 2126 BP.
AC AAK94505;
AT 06-NOV-2001 (first entry)
XX Human full-length cDNA, SEQ ID NO: 3356.
DE Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
KM Homo sapiens.
XX EP1130094-A2.
PN 05-SEP-2001.
XX 07-JUL-2000; 2000EP-0114089.
XX 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
XX (HELI-) HELIX RES INST.
PA Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Makamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI: 2001-524255/58.
DR P-PSDB: AAM93573.
XX 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX Claim 8: SEQ ID NO 3356; 1380bp + sequence listing; English.

XX The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesized by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a full length
CC human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.

XX Sequence 2126 BP; 501 A; 608 C; 520 G; 497 T; 0 other;
XX

Query Match 100.0%; Score 1269; DB 22; Length 2126;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGTGCACGAATGCTTTGACCTGATGCGCAGTCAAGAGCAGTGTAGATATGATGAA 60
DB 281 CAGTGCACGAATGCTTTGACCTGATGCGCAGTCAAGAGCAGTGTAGATATGATGAA 340

QY 61 TGCCGAACCATCCCGGAGGCTGCGGAGAGACATGATGTGTAAACCAATGCGGG 120
DB 341 TGCCGAACCATCCCGGAGGCTGCGGAGAGACATGATGTGTAAACCAATGCGGG 400
QY 121 TATTTATGCAATTCGCCGGACAAACCTGTGTATGAGAGGCCCTACTCGAACCCCTACTCG 180
DB 401 TATTTATGCAATTCGCCGGACAAACCTGTGTATGAGAGGCCCTACTCGAACCCCTACTCG 460
QY 181 ACCCCCTACTCAGAGTCCGTACCCAGCAGTGCCTCCACCTACCTAGCTCCAAATATCCC 240
DB 461 ACCCCCTACTCAGAGTCCGTACCCAGCAGTGCCTCCACCTACCTAGCTCCAAATATCCC 520
QY 241 AGCATCTCAAGCCCTCTTATATGCGCTTATATGATCCAGATGATGTAAGCAACCAATGT 300
DB 521 AGCATCTCAAGCCCTCTTATATGCGCTTATATGATCCAGATGATGTAAGCAACCAATGT 580
QY 301 GTGGATGTGAGCAGAGTGTGCAACAGATTCGCCACAGTCAACCCCAACCCAGATCTGCATC 360
DB 581 GTGGATGTGAGCAGAGTGTGCAACAGATTCGCCACAGTCAACCCCAACCCAGATCTGCATC 640
QY 361 AATACTGAAGGCGGTACACCTGCTCTGACCGAGGATATGGCTTGTGGAAGGCGAG 420
DB 641 AATACTGAAGGCGGTACACCTGCTCTGACCGAGGATATGGCTTGTGGAAGGCGAG 700
QY 421 TGCTTAGACATTTGATGATGTGCTATGCTTACTGCCAGCAGCTGTGCGAATGTTCT 480
DB 701 TGCTTAGACATTTGATGATGTGCTATGCTTACTGCCAGCAGCTGTGCGAATGTTCT 760
QY 481 GGATCCATTTCTTGTATGATGCAACCCCTGTTTACCTCAATAGATGGAAGTCTTGC 540
DB 761 GGATCCATTTCTTGTATGATGCAACCCCTGTTTACCTCAATAGATGGAAGTCTTGC 820
QY 541 CAAGATGTGAAGAGAGTGTGCAACCCCGAGAACCCCTGCGTGAACCTGCTCAACACCTAC 600
DB 821 CAAGATGTGAAGAGAGTGTGCAACCCCGAGAACCCCTGCGTGAACCTGCTCAACACCTAC 880
QY 601 GGCCTTTTCAATGTCGCGCTGTGACCCAGATATGAACCTGAGGAAGTGTGGAACCG 660
DB 881 GGCCTTTTCAATGTCGCGCTGTGACCCAGATATGAACCTGAGGAAGTGTGGAACCG 940
QY 661 AGTATATGAGAGAGTGTGCAACCCCGAGAACCCCTGCGTGAACCTGCTCAACACCTAC 720
DB 941 AGTATATGAGAGAGTGTGCAACCCCGAGAACCCCTGCGTGAACCTGCTCAACACCTAC 1000
QY 721 CCGGCAATACTTGTCTCTGCTGCTCCAGGCTACATCTGCTGATGACCAACCGAAGC 780
DB 1001 CCGGCAATACTTGTCTCTGCTGCTCCAGGCTACATCTGCTGATGACCAACCGAAGC 1060
QY 781 TGCCAAAGCATCAAGCAATGTGAGCAGAGAACCAACAGTCAACCTGCAAGCAGAGCTGC 840
DB 1061 TGCCAAAGCATCAAGCAATGTGAGCAGAGAACCAACAGTCAACCTGCAAGCAGAGCTGC 1120
QY 841 TACATTTTCAAGGGGCTTCAATATGATCGAACCCATCGCTGTGAGAGGCTTATCTG 900
DB 1121 TACATTTTCAAGGGGCTTCAATATGATCGAACCCATCGCTGTGAGAGGCTTATCTG 1180
QY 901 AGGATCAATGATTAACCGCTGTATGTGCTGCTGAGAACCTGTGCTGAGAGACCGCC 960
DB 1181 AGGATCAATGATTAACCGCTGTATGTGCTGCTGAGAACCTGTGCTGAGAGACCGCC 1240
QY 961 TTTACCATCTTGTACCGGAGATGAGAGTGTGAGAGAGCTCGTCCCTGCAATC 1020
DB 1241 TTTACCATCTTGTACCGGAGATGAGAGTGTGAGAGAGCTCGTCCCTGCAATC 1300
QY 1021 TTCCAAATGCAAGCAGACCGCTACCTGCGGCTTATATTTTCCAGATCAAAATGT 1080
DB 1301 TTCCAAATGCAAGCAGACCGCTACCTGCGGCTTATATTTTCCAGATCAAAATGT 1360
QY 1081 GGGATATGAGGCAAGAAATTTTACATGCGGCAACCGGCGCCATAGTGCACCCCTGCTG 1140
DB 1361 GGGATATGAGGCAAGAAATTTTACATGCGGCAACCGGCGCCATAGTGCACCCCTGCTG 1420

Human; ss; gene; extracellular protein-like protein; EGF-like; protein; epidermal growth factor; EGF; ATCC 97285; gene therapy; wound healing; neurologic trauma; acquired immunodeficiency syndrome; AIDS-related dementia; ocular disorder; kidney disorder; liver disorder; hair follicle growth promotion; burn; ulcer; corneal incision; corneal inflammation; neoplasm; psoriasis.

Homo sapiens.

Key: location/Qualifiers
CDS 211..1557
/tag- a
/product- "EBGF"
/note- "This CDS minus the stop codon (211-2554) is specifically claimed in claim 5"
sig-peptide 211..285
/tag- b
mat-peptide 286..1554
/tag- c
/label- Mature_EBGF
/note- "This region is specifically claimed in claim 4"

US2001051358-A1.

13-DEC-2001.

25-MAR-1999; 9905-0275805.

11-APR-1997; 9705-0839525.

10-APR-1996; 96WO-US05247.

(OLSEN) OLSEN H. S.
(LIH/), LI H.

Olsen HS, LI H;

WPI; 2002-121417/16.
P-PSDB; AAU75494.

New nucleic acid encoding human extracellular/epidermal growth factor, useful for diagnosis and treatment of e.g. Marfan syndrome and wounds, also related polypeptides

Claim 4; Fig 1; 22pp; English.

The invention relates to a novel polynucleotide which is at least 95% identical with a sequence (ATCC 97285) encoding mature human extracellular protein-like/epidermal growth factor (EGF)-like protein, EGF. Also included are the EGF EGF domains, a vector containing the polynucleotide, a host cell containing the vector, anti-EGF antibodies and antagonists of EGF. The polynucleotide is used for recombinant production of EGF, in gene therapy, as hybridisation probes, as antisense antagonists and for chromosome identification. The protein is used to treat patients who require EGF, to identify specific antagonists, used to treat conditions that require inhibition of EGF (e.g. vascular smooth muscle cell proliferation, Marfan syndrome, wound healing, neurological trauma, acquired immunodeficiency syndrome (AIDS)-related dementia, ocular disorders, kidney disorders, liver disorders, hair follicle growth promotion, burns, ulcers, corneal incisions, corneal inflammation, neoplasms and psoriasis), to raise specific antibodies and to characterise receptors. The present sequence encodes EGF.

Sequence 1720 BP; 396 A; 515 C; 434 G; 372 T; 3 other;

Query Match 100.0%; Score 1269; DB 24; Length 1720;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CAGTGCAGATGCTTGTACCTGGATGCGCAGTGCAGACAGTGTGTATGATGAA 60
|||||

286 CAGTGCAGATGCTTGTACCTGGATGCGCAGTGCAGACAGTGTGTATGATGAA 345
61 TGGCAACCATCCCGAGAGCGCTGGCGAGAGCATGATGTGTATACCAAAATGCGGG 120
|||||
346 TGGCAACCATCCCGAGAGCGCTGGCGAGAGCATGATGTGTATACCAAAATGCGGG 405
121 TATTATGATGATCCCGGAGCAAAACCTGTGTATGAGGCGCTTACTGCAACCCCTACTCG 180
406 TATTATGATGATCCCGGAGCAAAACCTGTGTATGAGGCGCTTACTGCAACCCCTACTCG 465
181 ACCCCCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
466 ACCCCCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 525
241 ACGAT 300
526 ACGAT 585
301 GTGAT 360
586 GTGAT 645
361 AATAGTGAAGGCGGCTACACCTGCTCTGACACGAGATATGCTTGTGAGAGCCAG 420
646 AATAGTGAAGGCGGCTACACCTGCTCTGACACGAGATATGCTTGTGAGAGCCAG 705
421 TGTATGAT 480
706 TGTATGAT 765
481 GGAT 540
766 GGAT 825
541 CAAGAT 600
826 CAAGAT 885
601 GGCTCTTATCATCTGCGCTGTGACCCAGATATGAACTGAGGAAAGATGAGCGTTCAATGTC 660
886 GGCTCTTATCATCTGCGCTGTGACCCAGATATGAACTGAGGAAAGATGAGCGTTCAATGTC 945
661 AGTATATGAGAGAGTGCAGCTTCTGTAGTCTCTGAGTCTCTGAGTATGAGTGTGAACAG 720
946 AGTATATGAGAGAGTGCAGCTTCTGTAGTCTCTGAGTCTCTGAGTATGAGTGTGAACAG 1005
721 CCGGACATATCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
1006 CCGGACATATCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1065
781 TGGCAAGCATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
1066 TGGCAAGCATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1125
841 TACAATTTACAAGGGGCTTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
1126 TACAATTTACAAGGGGCTTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1185
901 AGAT 960
1186 AGAT 1245
961 TTTACCATCTTGTACCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
1246 TTTACCATCTTGTACCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1305
1021 TTCCAAATGCAAGCCAGACCCGCTACCTGGGGCTATTTACATTTTCCAGATCAATCT 1080
1306 TTCCAAATGCAAGCCAGACCCGCTACCTGGGGCTATTTACATTTTCCAGATCAATCT 1365
1081 GGGATGATGAGGCAAGATTTTATCATGCGGCAAAAGGGGCCCATAGTCCACCCCTGGTG 1140
1366 GGGATGATGAGGCAAGATTTTATCATGCGGCAAAAGGGGCCCATAGTCCACCCCTGGTG 1425

PD 22-OCT-1998.
 XX 11-APR-1997; 97MO-US06020.
 XX 11-APR-1997; 97MO-US06020.
 PR (HUMA-) HUMAN GENOME SCI INC.
 XX
 PA LI H, Olsen HS;
 XX
 PI MPI: 1998-568728/48.
 XX P-PSDB; AAM/9739.
 DR
 DR
 PT New isolated extracellular/epidermal growth factor - used for
 PT regulating vascular smooth muscle cell proliferation, e.g. for
 PT enhancing neurological functions or treating neoplasia and other
 PT disorders.

PS Claim 1a; Fig 1A-D; 62bp; English.

CC This sequence encodes a novel human extracellular/epidermal growth
 CC factor-like protein, EGF. This protein can be used to regulate
 CC vascular smooth muscle cell proliferation and for restoration or
 CC enhancement of neurological functions diminished as a result of other
 CC damaging pathologies such as AIDS dementia. The protein can also be used
 CC to treat senile dementia, ocular disorders such as corneal inflammation,
 CC for targeting tumor cells, for treating kidney disorders, for liver
 CC regeneration or treating liver dysfunction, for treating wounds including
 CC all cutaneous wounds, corneal wounds, and injuries to the
 CC epithelial-lined hollow organs of the body or resulting from trauma such
 CC as burns, abrasions and cuts as well as from surgical procedures such as
 CC surgical incisions and skin grafting. The polypeptides can also be used
 CC for treating chronic conditions, such as chronic ulcers, diabetic ulcers,
 CC other non-healing (trophic) conditions, to treat Marfan syndrome, to
 CC promote hair follicular development, to stimulate growth and
 CC differentiation of various epidermal and epithelial cells in vivo and in
 CC vitro and to stimulate embryogenesis. Antagonists to EGF can be used to
 CC treat neoplasia such as cancers or tumours, skin disorders such as
 CC psoriasis or corneal inflammation. The products can also be used for
 CC identifying EGF receptors, detection, diagnosis and drug screening.

XX Sequence 1720 BP; 396 A; 515 C; 434 G; 372 T; 3 other;

Query Match 100.0%; Score 1269; DB 19; Length 1720;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 CAGTGCACGAGTGGCTTTGACCTGGATGCCAGTCAGAGAGGTGTTAGATTGATGA 60
 DB 286 CAGTGCACGAGTGGCTTTGACCTGGATGCCAGTCAGAGAGGTGTTAGATTGATGA 345
 QY 61 TGGCCGACCATCCCGAGGCGCTCCGAGAGACATGATGTTTAAACCAAAATGGCGG 120
 DB 346 TCCCGAACCATCCCGAGGCGCTCCGAGAGACATGATGTTTAAACCAAAATGGCGG 405
 QY 121 TATTTATGATTCCTCCGAGAACCCGTGTATGAGAGGCGCTACTGAAACCCCTACTCG 180
 DB 406 TATTTATGATTCCTCCGAGAACCCGTGTATGAGAGGCGCTACTGAAACCCCTACTCG 465
 QY 181 ACCCCCTACTCAGGTCCGTAACCAAGAGTGGCCCAACCACTCTCAGTCAAACTATCC 240
 DB 466 ACCCCCTACTCAGGTCCGTAACCAAGAGTGGCCCAACCACTCTCAGTCAAACTATCC 525
 QY 241 ACGATCTCCAGGCTCTTATATGCCGTTTGGATACAGATGATGAAGAACAACATGT 300
 DB 536 ACGATCTCCAGGCTCTTATATGCCGTTTGGATACAGATGATGAAGAACAACATGT 585
 QY 301 GTGGATGTGACAGAGTGTGAACAAGATTCCACAGTGAACCCCAACCAATCTGCATC 360
 DB 586 GTGGATGTGACAGAGTGTGAACAAGATTCCACAGTGAACCCCAACCAATCTGCATC 645
 QY 361 AATAGTGAAGCGGCTACACTGCTCTGACAGCAGATGATTTGGCTTCTGGAAGCGCAG 420

DB 646 AATAGTGAAGCGGCTACACTGCTCTGACAGCAGATGATTTGGCTTCTGGAAGCGCAG 705
 QY 421 TGGTTAGACATTTGATGATGATGCTGATGATGCTGACGAGCTGTCGGAATGTGCT 480
 DB 706 TGGTTAGACATTTGATGATGATGCTGATGATGCTGACGAGCTGTCGGAATGTGCT 765
 QY 481 GGATCTTATCTTGTATACATGCAACCTGTTTTTACCTCAATGAGATGGAAGTCTTGC 540
 DB 766 GGATCTTATCTTGTATACATGCAACCTGTTTTTACCTCAATGAGATGGAAGTCTTGC 825
 QY 541 CAAGATGTGAACAGATGTGCCACCGAGAACCCCTGCGTGTCAAACTGCTCAACCTTAC 600
 DB 826 CAAGATGTGAACAGATGTGCCACCGAGAACCCCTGCGTGTCAAACTGCTCAACCTTAC 885
 QY 601 GGCTCTTTCATCTGCGGCTGTGACCCAGGATATGAACTTGGAAGATGGGCTTATGCG 660
 DB 886 GGCTCTTTCATCTGCGGCTGTGACCCAGGATATGAACTTGGAAGATGGGCTTATGCG 945
 QY 661 AGTATATGAGACAGTGCACCTTCTGTAGTTCTCTGCGCAACATGAGTGTGAACGAG 720
 DB 946 AGTATATGAGACAGTGCACCTTCTGTAGTTCTCTGCGCAACATGAGTGTGAACGAG 1005
 QY 721 CCCGACATATCTTCTCTCTGCTGCTGCGCTGACGCTACATCTGCTGATGACACCGAAGC 780
 DB 1006 CCCGACATATCTTCTCTCTGCTGCTGCGCTGACGCTACATCTGCTGATGACACCGAAGC 1065
 QY 781 TGGCAAGACATCAACGAATGAGACAGAGAACCAACGTGCAACCTGACAGAGAGTGC 840
 DB 1066 TGGCAAGACATCAACGAATGAGACAGAGAACCAACGTGCAACCTGACAGAGAGTGC 1125
 QY 841 TACAATTTACAGGGGGCTTCAATGATCGACACCCATCCGCTGTGAGAGGCTTATGTC 900
 DB 1126 TACAATTTACAGGGGGCTTCAATGATCGACACCCATCCGCTGTGAGAGGCTTATGTC 1185
 QY 901 AGGATAGTATTAACCGCTGTATGTCTGCTGAGAAACCTGCTGACAGACACGCC 960
 DB 1186 AGGATAGTATTAACCGCTGTATGTCTGCTGAGAAACCTGCTGACAGACACGCC 1245
 QY 961 TTATACATCTTGTACCGGAGCATGAGAGTGTGTCAGAGCTGCTCCGCTGACATC 1020
 DB 1246 TTATACATCTTGTACCGGAGCATGAGAGTGTGTCAGAGCTGCTCCGCTGACATC 1305
 QY 1021 TTCCAATGACACGACACGACCCGCTACCTGGGGCTTATGATTTTCCAGATCAAACT 1080
 DB 1306 TTCCAATGACACGACACGACCCGCTACCTGGGGCTTATGATTTTCCAGATCAAACT 1365
 QY 1081 GGGATATGAGGGGACAGAAATTTTACATGCGGCAACCGGCCCATATCAGTCCAGCTG 1140
 DB 1366 GGGATATGAGGGGACAGAAATTTTACATGCGGCAACCGGCCCATATCAGTCCAGCTG 1425
 QY 1141 ATGACAGGCGCCCAATCAAGAGGCGCCGGAATTCAGCTGGAATGATGATCACTGTC 1200
 DB 1426 ATGACAGGCGCCCAATCAAGAGGCGCCGGAATTCAGCTGGAATGATGATCACTGTC 1485
 QY 1201 AACACGTCTATCACTTCAAGAGGAGCTGCTGATTCGCACTGCGGATATATGTGTCGAG 1260
 DB 1486 AACACGTCTATCACTTCAAGAGGAGCTGCTGATTCGCACTGCGGATATATGTGTCGAG 1545
 QY 1261 TACCCATTC 1269
 DB 1546 TACCCATTC 1554

RESULT 3

ABK13627
 ID ABK13627 standard; cDNA; 1720 BP.

XX ABK13627;

AC 23-APR-2002 (first entry)

XX Human cDNA encoding extracellular protein-1like/EGF-like protein, EGF.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 08:40:41 ; Search time 282.77 Seconds
(without alignments)
10106.379 Million cell updates/sec

Title: US-09-674-379a-15
Perfect score: 1269
Sequence: 1 cagtcacgaatgccttga.....atgtcgcagcaccaccac 1269

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N.Geneseq_101002.*

1:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1269	100.0	1347	24	ABL52528
2	1269	100.0	1720	19	AAV62432
3	1269	100.0	1720	24	ABK13627
4	1269	100.0	2126	22	AAK94505
5	1269	100.0	2328	21	AAZ40027
6	1269	100.0	2362	20	AAK07567
7	1269	100.0	2362	21	AAZ39692
8	1269	100.0	2550	20	AAK05359
9	1267.4	99.9	2509	20	AAK37670

10	1267.4	99.9	2609	22	AA546128	Human DNA encoding
11	1267.4	99.9	2609	22	AA86968	Nucleotide sequenc
12	1267.2	95.1	1228	18	AA89380	Human extracellular
13	1056.2	83.2	1269	21	AA339385	Smooth muscle prol
14	1056.2	83.2	1269	21	AA339386	Smooth muscle prol
15	1056.2	83.2	1344	21	AA339383	Smooth muscle prol
16	1056.2	83.2	1383	21	AA339386	Smooth muscle prol
17	1056.2	83.2	2233	21	AA339384	Smooth muscle prol
18	1056.2	83.2	2233	21	AA400029	Full length mouse
19	1056.2	83.2	2429	21	AA339387	Smooth muscle prol
20	1056.2	83.2	2429	21	AA400031	Full length mouse
21	843.8	66.5	1254	21	AA61672	CDNA encoding rat
22	843.8	66.5	1254	22	AA69605	Skin cell CDNA, SE
23	841.8	66.3	1260	24	AB34757	Rat CDNA isolated
24	776.4	61.2	778	20	AA337673	Human PRO protein
25	721.4	56.8	853	22	AA91596	Human CDNA 5'-end
26	721.4	56.8	853	22	AA93169	Human CDNA clone r
27	721.4	56.8	1538	22	AA94129	Human CDNA encoding
28	576.8	45.5	680	21	AA44245	Human secreted exp
29	374.4	29.5	377	20	AA337686	Human EST sequence
30	371.2	29.3	1480	21	AA14608	CDNA encoding a hu
31	371.2	29.3	1525	18	AA188974	Human extracellular
32	371.2	29.3	1531	20	AA60351	Extracellular prot
33	371.2	29.3	1707	22	AAH13771	Human CDNA sequenc
34	371.2	29.3	1874	21	AA58583	Human PRO226 prote
35	371.2	29.3	1875	22	AA521402	Human CDNA sequenc
36	371.2	29.3	1875	22	AA40256	CDNA encoding huma
37	371.2	29.3	1994	21	AA18229	Lung cancer associ
38	371.2	29.3	2018	21	AA39800	Human SI-5 ECKP-11
39	347.2	27.4	1513	21	AA14603	CDNA encoding a p5
40	337.6	26.6	1358	21	AA14616	CDNA encoding the
41	326	25.7	354	20	AA337683	Human EST sequence
42	316.6	24.9	1018	21	AA61786	CDNA encoding huma
43	316.6	24.9	1018	22	AA99719	Skin cell CDNA, SE
44	316.6	24.9	1018	22	AB34871	Murine CDNA isolat
45	315	24.8	670	22	AA92195	Human CDNA 5'-end

ALIGNMENTS

RESULT 1					
ABL52528	ABL52528	standard;	CDNA;	1347	BP.
XX	XX				
AC	ABL52528;				
XX	XX				
DT	19-JUL-2002	(first entry)			
XX	XX				
DE	Human UP50	CDNA sequence SEQ ID NO:1.			
XX	XX				
KW	Human; UP50;	artificial vascular graft; cellular adherence factor;			
KW	cell proliferating	growth factor; DANCE; epidermal growth factor; EGF;			
XX	developmental	arteries and neural crest EGF-like protein; gene; ss.			
OS	Homo sapiens.				
XX	XX				
PN	WO200207646-A2.				
PD	31-JAN-2002.				
XX	XX				
PF	20-JUL-2001;	2001MO-IL00670.			
XX	XX				
PR	20-JUL-2000;	2000US-0620227.			
XX	XX				
PA	(MGVS-) MGVS LTD.				
XX	XX				
PI	Flugelman MY, Preis M, Gluzman Z, Koren B, Weisz A, Cohen T;				
XX	XX				
DR	WPI, 2002-361632/39.				
XX	XX				
PT	Artificial vascular graft for therapy of vascular diseases, has				
XX	synthetic tubular element having a surface coated with endothelial,				

Sat Jul 5 15:31:34 2003

us-09-674-379a-15.rnpb

Page 16

Search completed: July 3, 2003, 17:53:39
Job time : 207.749 secs

OY 961 TTACCATCTGTACCGGACATGACGTGTGTGACGACCTCCGTTCCCGTGACATC 1020
|||||
Db 1485 TTACCATCTGTGTACCGGACATGACGTGTGTGACGACCTCCGTTCCCGTGACATC 1544
|||||
OY 1021 TTCCAAATGCAAGCAGCAGCCGCTACCTGTGGGGCTATTACATTTTCCAGATCAATCT 1080
|||||
Db 1545 TTCCAAATGCAAGCAGCAGCCGCTACCTGTGGGGCTATTACATTTTCCAGATCAATCT 1604
|||||
OY 1081 GGGATATGAGGAGAGAAATTTTACATGCGGCAAAAGGGCCCATGATGCGACCTGTGTG 1140
|||||
Db 1605 GGGATATGAGGAGAGAAATTTTACATGCGGCAAAAGGGCCCATGATGCGACCTGTGTG 1664
|||||
OY 1141 ATGACACGCCCATCAAAAGGGCCCGGAAATCCAGCTGGACTTGGAAATGATCACTGTC 1200
|||||
Db 1665 ATGACACGCCCATCAAAAGGGCCCGGAAATCCAGCTGGACTTGGAAATGATCACTGTC 1724
|||||
OY 1201 AACACTGTATCACTTCAAGAGCAGCTCCGTATCCGACTCGGATATATGTGTGCGAG 1260
|||||
Db 1725 AACACTGTATCACTTCAAGAGCAGCTCCGTATCCGACTCGGATATATGTGTGCGAG 1784
|||||
OY 1261 TACCATTC 1269
|||||
Db 1785 TACCATTC 1793
|||||

RESULT 15

US-10-180-552-407
; Sequence 407, Application US/10180552
; Publication No. US20030022300A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Matanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C153
; CURRENT APPLICATION NUMBER: US/10/180.552
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 407
; LENGTH: 2609
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-180-552-407

Query Match 99.98; Score 1267.4; DB 9; Length 2609;
Best Local Similarity 99.98; Pred. No. 0;
Matches 1268; Conservative 0; Mismatches 1; Incls 0; Gaps 0;

OY 1 CAGTGCACGAATGGCTTTGACCTGATCGCAGTCAGACAGTGTGTTAGATTTGATGAA 60
|||||
Db 525 CAGTGCACGAATGGCTTTGACCTGATCGCAGTCAGACAGTGTGTTAGATTTGATGAA 584
|||||
OY 61 TGGCGAACATCCCGGAGGCTGCGGAGAGACATGATGTGTTAAACCAAAATGGCGG 120
|||||
Db 585 TGGCGAACATCCCGGAGGCTGCGGAGAGACATGATGTGTTAAACCAAAATGGCGG 644
|||||
OY 121 TATTTATGATTCGCCGAGAAACCTGTGTATGAGAGGCCCTACTCGAACCCCTACTCG 180
|||||
Db 645 TATTTATGATTCGCCGAGAAACCTGTGTATGAGAGGCCCTACTCGAACCCCTACTCG 704
|||||
OY 181 ACCCGCTACTCAGTCCGTATCCGAGAGAGCTGCCACCATCTCTAGCTCCAAATCTATCC 240
|||||
Db 705 ACCCGCTACTCAGTCCGTATCCGAGAGAGCTGCCACCATCTCTAGCTCCAAATCTATCC 764
|||||

OY 241 AGCATCTCCAGGCGCTTATATAGCCGCTTTGGATACAGATGATGAAAGCAACCAATGT 300
|||||
Db 765 AGCATCTCCAGGCGCTTATATATGCGGCTTTGGATACAGATGATGAAAGCAACCAATGT 824
|||||
OY 301 GTGATGTGAGAGAGTGTGCAACAGATTTCCACAGAGTGCACCCCAACCCAGATCTGCATC 360
|||||
Db 825 GTGATGTGAGAGAGTGTGCAACAGATTTCCACAGAGTGCACCCCAACCCAGATCTGCATC 884
|||||
OY 361 AATAGTGAAGCGGGTACACCTGCTCCGACCGGAGGATTTGGCTCTGGAAGGCCAG 420
|||||
Db 885 AATAGTGAAGCGGGTACACCTGCTCCGACCGGAGGATTTGGCTCTGGAAGGCCAG 944
|||||
OY 421 TGTATGACATTTGATGATGATGCTGATGTTACTGTCACAGACCTCTGTGCAATGTCTCT 480
|||||
Db 945 TGTATGACATTTGATGATGATGCTGATGTTACTGTCACAGACCTCTGTGCAATGTCTCT 1004
|||||
OY 481 GGATCCATTTCTTTTACATGCAACCTGTGTTTACCTTCATGAGGATGGAAGTCTTGC 540
|||||
Db 1005 GGATCCATTTCTTTTACATGCAACCTGTGTTTACCTTCATGAGGATGGAAGTCTTGC 1064
|||||
OY 541 CAAGATGTGAACGAGTGTGACCGAGAGAACCCCTGTCAAACCTGGGTCAACACCTTAC 600
|||||
Db 1065 CAAGATGTGAACGAGTGTGACCGAGAGAACCCCTGTCAAACCTGGGTCAACACCTTAC 1124
|||||
OY 601 GCGCTTTTCAATCTGCGCGTGTGACCCAGAGATATGAATTTGAGAAATGGCGTTTATTC 660
|||||
Db 1125 GCGCTTTTCAATCTGCGCGTGTGACCCAGAGATATGAATTTGAGAAATGGCGTTTATTC 1184
|||||
OY 661 AGTATGTGAGAGAGTGTGACCTGCTGAGTGTCTGTCACCAACATGATGTGTGAACAG 720
|||||
Db 1185 AGTATGTGAGAGAGTGTGACCTGCTGAGTGTCTGTCACCAACATGATGTGTGAACAG 1244
|||||
OY 721 CCGGACATATCTTCTGCTGCTGCTCCAGAGTACATCTGTGTGATGACAACCGAAGC 780
|||||
Db 1245 CCGGACATATCTTCTGCTGCTGCTCCAGAGTACATCTGTGTGATGACAACCGAAGC 1304
|||||
OY 781 TGGCAGACATCAACGAATGTGACACAGAGAACACACAGCTGACAGACGTGC 840
|||||
Db 1305 TGGCAGACATCAACGAATGTGACACAGAGAACACACAGCTGACAGACGTGC 1364
|||||
OY 841 TACAATTTACAAGGGGCTTCAATGATGATGACACCCATCGCTGTGAGAGCCTTATCTG 900
|||||
Db 1365 TACAATTTACAAGGGGCTTCAATGATGATGACACCCATCGCTGTGAGAGCCTTATCTG 1424
|||||
OY 901 AGGATCAGTATAACCGCTGTATGTCTGCTGAGAAACCTGTGTCAGAGACAGCC 960
|||||
Db 1425 AGGATCAGTATAACCGCTGTATGTCTGCTGAGAAACCTGTGTCAGAGACAGCC 1484
|||||
OY 961 TTTACCATCTTTGACCGGAGACATGAGAGTGTGTCAGAGACCTCCGTTCCGTCATTC 1020
|||||
Db 1485 TTTACCATCTTTGACCGGAGACATGAGAGTGTGTCAGAGACCTCCGTTCCGTCATTC 1544
|||||
OY 1021 TTCCAAATGCAAGCAGCAGCCGCTACCTGTGGGGCTATTACATTTTCCAGATCAATCT 1080
|||||
Db 1545 TTCCAAATGCAAGCAGCAGCCGCTACCTGTGGGGCTATTACATTTTCCAGATCAATCT 1604
|||||
OY 1081 GGGATATGAGGAGAGAAATTTTACATGCGGCAAAAGGGCCCATGATGCGACCTGTGTG 1140
|||||
Db 1605 GGGATATGAGGAGAGAAATTTTACATGCGGCAAAAGGGCCCATGATGCGACCTGTGTG 1664
|||||
OY 1141 ATGACACGCCCATCAAAAGGGCCCGGAAATCCAGCTGGACTTGGAAATGATCACTGTC 1200
|||||
Db 1665 ATGACACGCCCATCAAAAGGGCCCGGAAATCCAGCTGGACTTGGAAATGATCACTGTC 1724
|||||
OY 1201 AACACTGTATCACTTCAAGAGCAGCTCCGTATCCGACTCGGATATATGTGTGCGAG 1260
|||||
Db 1725 AACACTGTATCACTTCAAGAGCAGCTCCGTATCCGACTCGGATATATGTGTGCGAG 1784
|||||
OY 1261 TACCATTC 1269
|||||
Db 1785 TACCATTC 1793
|||||

241 ACAGTCACAGGCTCTTATATGCGCTTTGGATACAGATGATGAAGCAACCAATGT 300
|||||
765 ACAGTCACAGGCTCTTATATGCGCTTTGGATACAGATGATGAAGCAACCAATGT 824
|||||
301 GTGGATGTGACGAGTGTGCAACAGATTCCACAGTGCACACCCACAGTTCGCAATC 360
|||||
825 GTGGATGTGACGAGTGTGCAACAGATTCCACAGTGCACACCCACAGTTCGCAATC 884
|||||
361 AATACGAAAGGCGGGTACACCTGCTCTGCAACGAGATATGGCTTCTGAAAGGCCAG 420
|||||
885 AATACGAAAGGCGGGTACACCTGCTCTGCAACGAGATATGGCTTCTGAAAGGCCAG 944
|||||
421 TGCTTAGACATGTAGTAATGCGCTATGATGCTAGCGACAGCTCTGCGCAATGTCTCT 480
|||||
945 TGCTTAGACATGTAGTAATGCGCTATGATGCTAGCGACAGCTCTGCGCAATGTCTCT 1004
|||||
481 GGATCTTATCTTGTACATGCAACCCCTGATTTTACCTCAATGAGATGAAAGCTTGC 540
|||||
1005 GGATCTTATCTTGTACATGCAACCCCTGATTTTACCTCAATGAGATGAAAGCTTGC 1064
|||||
541 CAAGATGTGAACGAGTGTGCAACGAGATACCCCTGCTGCAACCTGCTGCAACCTTAC 600
|||||
1065 CAAGATGTGAACGAGTGTGCAACGAGATACCCCTGCTGCAACCTGCTGCAACCTTAC 1124
|||||
601 GGCTCTTATCTTGTACATGCAACCCCTGATTTTACCTCAATGAGATGAAAGCTTGC 660
|||||
1125 GGCTCTTATCTTGTACATGCAACCCCTGATTTTACCTCAATGAGATGAAAGCTTGC 1184
|||||
661 AGTGATGTGACGAGTGTGCAACGAGATACCCCTGCTGCAACCTGCTGCAACCTTAC 720
|||||
1185 AGTGATGTGACGAGTGTGCAACGAGATACCCCTGCTGCAACCTGCTGCAACCTTAC 1244
|||||
721 CCCGGACATCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
|||||
1245 CCCGGACATCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1304
|||||
781 TGCCAAAGACATCAACGAATGTGACAGAGAACCAACAGTGAACCTGCAACAGCTGC 840
|||||
1305 TGCCAAAGACATCAACGAATGTGACAGAGAACCAACAGTGAACCTGCAACAGCTGC 1364
|||||
841 TACAAATTTCAAGGGGGCTTCAAAATGCATGCAACCCATCCGCTGTGAGAGCTTATCTG 900
|||||
1365 TACAAATTTCAAGGGGGCTTCAAAATGCATGCAACCCATCCGCTGTGAGAGCTTATCTG 1424
|||||
901 AGGATCAGTATACCGCTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
|||||
1425 AGGATCAGTATACCGCTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1484
|||||
961 TTTTACATCTTGTACCGGACATGAGACGTGTGTGAGAGCTCCGTTCCCGCTGACATC 1020
|||||
1485 TTTTACATCTTGTACCGGACATGAGACGTGTGTGAGAGCTCCGTTCCCGCTGACATC 1544
|||||
1021 TTTCCAAATGCAAGCAGCAGCGGCTACCTGGGGCTTATTAATTTTCCAGATCAATCT 1080
|||||
1545 TTTCCAAATGCAAGCAGCAGCGGCTACCTGGGGCTTATTAATTTTCCAGATCAATCT 1604
|||||
1081 GGGATGAGAGGAGAGATTTTACATGCGGCAAAAGGGCCCTCATGCTGCAACCTGCTG 1140
|||||
1605 GGGATGAGAGGAGAGATTTTACATGCGGCAAAAGGGCCCTCATGCTGCAACCTGCTG 1664
|||||
1141 ATGACACCCCTCATAAAGGGCCCGGGGAATTCACAGTGGACTTGGAAATGATCACTGTC 1200
|||||
1665 ATGACACCCCTCATAAAGGGCCCGGGGAATTCACAGTGGACTTGGAAATGATCACTGTC 1724
|||||
1201 AACACTGTCACTAATTTAGAGAGAGCTCCGCTGATCCGCTGCGATATATGTGTGCGAG 1260
|||||
1725 AACACTGTCACTAATTTAGAGAGAGCTCCGCTGATCCGCTGCGATATATGTGTGCGAG 1784
|||||
1261 TACCATTC 1269
|||||
1785 TACCATTC 1793

RESULT 13
US-10-176-757-407
; Sequence 407, Application US/10176757
; Publication No. US20030022297A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C86
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 407
; LENGTH: 2609
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-757-407
Query Match 99.9%; Score 1267.4; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 CAGTCACGAATGAGTGTGACCTGATCGCAGTGCAGAGTGTAGATATGATGA 60
DB CAGTCACGAATGAGTGTGACCTGATCGCAGTGCAGAGTGTAGATATGATGA 584
61 TGCCGAACATCCCGAGGCGCTGCGAGAGACATGATGTGTAAACCAATATGCGGG 120
DB TGCCGAACATCCCGAGGCGCTGCGAGAGACATGATGTGTAAACCAATATGCGGG 644
585 TGCCGAACATCCCGAGGCGCTGCGAGAGACATGATGTGTAAACCAATATGCGGG 644
121 TATTATGATTCCTCCCGGCAAAACCTGTGTGAGAGGCGCTTACGCAACCTTACTG 180
DB TATTATGATTCCTCCCGGCAAAACCTGTGTGAGAGGCGCTTACGCAACCTTACTG 645
645 TATTATGATTCCTCCCGGCAAAACCTGTGTGAGAGGCGCTTACGCAACCTTACTG 704
181 ACCCCCTACTGAGGCTTATATGCGCTTTGGATACAGATGATGAAGCAACCAATGT 300
DB ACCCCCTACTGAGGCTTATATGCGCTTTGGATACAGATGATGAAGCAACCAATGT 705
705 ACCCCCTACTGAGGCTTATATGCGCTTTGGATACAGATGATGAAGCAACCAATGT 764
241 ACGATTCGACAGGCTCTTATATGCGCTTTGGATACAGATGATGAAGCAACCAATGT 300
DB ACGATTCGACAGGCTCTTATATGCGCTTTGGATACAGATGATGAAGCAACCAATGT 765
765 ACGATTCGACAGGCTCTTATATGCGCTTTGGATACAGATGATGAAGCAACCAATGT 824
301 GTGATGTGAGAGTGTGCAACAGATTCGCCACAGTGAACCCACCAAGATTCGATC 360
DB GTGATGTGAGAGTGTGCAACAGATTCGCCACAGTGAACCCACCAAGATTCGATC 825
825 GTGATGTGAGAGTGTGCAACAGATTCGCCACAGTGAACCCACCAAGATTCGATC 884
361 AATACGAAAGGCGGGTACACCTGCTCTGCAACGAGATATGGCTTCTGAAAGGCCAG 420
DB AATACGAAAGGCGGGTACACCTGCTCTGCAACGAGATATGGCTTCTGAAAGGCCAG 885
885 AATACGAAAGGCGGGTACACCTGCTCTGCAACGAGATATGGCTTCTGAAAGGCCAG 944
421 TGCTTAGACATGTAGTAATGCGCTATGATGCTAGCGACAGCTCTGCGCAATGTCTCT 480
DB TGCTTAGACATGTAGTAATGCGCTATGATGCTAGCGACAGCTCTGCGCAATGTCTCT 945
945 TGCTTAGACATGTAGTAATGCGCTATGATGCTAGCGACAGCTCTGCGCAATGTCTCT 1004
481 GGATCTTATCTTGTACATGCAACCCCTGATTTTACCTCAATGAGATGAAAGCTTGC 540
DB GGATCTTATCTTGTACATGCAACCCCTGATTTTACCTCAATGAGATGAAAGCTTGC 1005
1005 GGATCTTATCTTGTACATGCAACCCCTGATTTTACCTCAATGAGATGAAAGCTTGC 1064
541 CAAGATGTGAACGAGTGTGCAACGAGATACCCCTGCTGCAACCTGCTGCAACCTTAC 600
DB CAAGATGTGAACGAGTGTGCAACGAGATACCCCTGCTGCAACCTGCTGCAACCTTAC 1065

TYPE: DNA
ORGANISM: Homo Sapien
US-10-175-752-407

Query Match 99.9% Score 1267.4; DB 9; Length 2609;
Best Local Similarity 99.9% Pred. No. 0;
Matches 1268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 CAGTGCAGCAATGGCTTGTGACCTGATGCGCAGTCCAGACAGTGTGATATGATGAA 60
DB 525 CAGTGCAGCAATGGCTTGTGACCTGATGCGCAGTCCAGACAGTGTGATATGATGAA 584
QY 61 TGGCGAACCATCCCGAGGCTGCGGAGAGACATGATGTGTAAACCAAAATGCGGAG 120
DB 585 TGGCGAACCATCCCGAGGCTGCGGAGAGACATGATGTGTAAACCAAAATGCGGAG 644
QY 121 TATTATGCAATTCCTCCGAGAAACCTGTGTATGAGAGGCTTACTGAAACCTTACTG 180
DB 645 TATTATGCAATTCCTCCGAGAAACCTGTGTATGAGAGGCTTACTGAAACCTTACTG 704
QY 181 ACCCTTACTGAGCTGCGTACCGAGAGCTGCGCAGACCTGCTGACCTCAAACTATGCC 240
DB 705 ACCCTTACTGAGCTGCGTACCGAGAGCTGCGCAGACCTGCTGACCTCAAACTATGCC 764
QY 241 ACGATCTCAGGCTCTTATATGCGGCTTGTGATACAGATGATGAAAGCAACCAATGT 300
DB 765 ACGATCTCAGGCTCTTATATGCGGCTTGTGATACAGATGATGAAAGCAACCAATGT 824
QY 301 GTGGATGTGACAGATGTGCACAGATTTCCACAGTGCACACCCCACTGATCTGCATC 360
DB 825 GTGGATGTGACAGATGTGCACAGATTTCCACAGTGCACACCCCACTGATCTGCATC 884
QY 361 AATACTGAAGGCGGCTGACCTGCTGACAGAGATGATGATGATGATGATGATGATG 420
DB 885 AATACTGAAGGCGGCTGACCTGCTGACAGAGATGATGATGATGATGATGATGATG 944
QY 421 TGGTATGACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
DB 945 TGGTATGACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1004
QY 481 GGAATCTATTTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 540
DB 1005 GGAATCTATTTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1064
QY 541 CAAGATGTGAAGAGTGTGCACAGAGAACCCGCGGGAACCTGCGTCAACACCTATC 600
DB 1065 CAAGATGTGAAGAGTGTGCACAGAGAACCCGCGGGAACCTGCGTCAACACCTATC 1124
QY 601 GGGCTCTTCAATCTGCGGCTGTGACCAAGATGATGATGATGATGATGATGATG 660
DB 1125 GGGCTCTTCAATCTGCGGCTGTGACCAAGATGATGATGATGATGATGATGATG 1184
QY 661 AGTATGATGACAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 720
DB 1185 AGTATGATGACAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1244
QY 721 CCCGGACATATCTTGTGCTGCGCTGCGAGCTACATCTGCGAGATGATGATGATG 780
DB 1245 CCCGGACATATCTTGTGCTGCGCTGCGAGCTACATCTGCGAGATGATGATGATG 1304
QY 781 TGGCAAGACATCAAGATGTGAGCAGAGAACCAACAGTGCACACCTGACAGAGCTGC 840
DB 1305 TGGCAAGACATCAAGATGTGAGCAGAGAACCAACAGTGCACACCTGACAGAGCTGC 1364
QY 841 TACAATTTACAAAGGGGCTTCAATGATGATGATGATGATGATGATGATGATGATG 900
DB 1365 TACAATTTACAAAGGGGCTTCAATGATGATGATGATGATGATGATGATGATGATG 1424
QY 901 AGGATAGTATGAACGCTGATGTGCTGATGATGATGATGATGATGATGATGATGATG 960
DB 1425 AGGATAGTATGAACGCTGATGTGCTGATGATGATGATGATGATGATGATGATGATG 1484
QY 961 TTATACATCTTGTACCGGAGCATGAGAGTGTGTACAGAGCTCCGTTCCGCTGACATC 1020

```

```

DB 1485 TTATACATCTTGTACCGGAGCATGAGAGTGTGTACAGAGCTCCGTTCCGCTGACATC 1544
QY 1021 TTCCAAATGCAAGCAGACAGACCCCTTACCCTGGGCTTATATATTTCCAGATCAAAAT 1080
DB 1545 TTCCAAATGCAAGCAGACAGACCCCTTACCCTGGGCTTATATATTTCCAGATCAAAAT 1604
QY 1081 GGAATAGAGGAGAGATTTTATATGATGATGATGATGATGATGATGATGATGATGATG 1140
DB 1605 GGAATAGAGGAGAGATTTTATATGATGATGATGATGATGATGATGATGATGATGATG 1664
QY 1141 ATGACAGCCCATCAAAAGGCGCCCGGAAATTCAGCTGATGATGATGATGATGATGATG 1200
DB 1665 ATGACAGCCCATCAAAAGGCGCCCGGAAATTCAGCTGATGATGATGATGATGATGATG 1724
QY 1201 AACAGTGTCAATCACTTCAAGAGAGCTGCGTGTATGATGATGATGATGATGATGATG 1260
DB 1725 AACAGTGTCAATCACTTCAAGAGAGCTGCGTGTATGATGATGATGATGATGATGATG 1784
QY 1261 TACCATTC 1269
DB 1785 TACCATTC 1793

```

RESULT 12
US-10-176-482-407
Sequence 407, Application US/10176482
Publication No. US20030022296A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zhen
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C70
CURRENT APPLICATION NUMBER: US/10/176,482
PRIORITY FILING DATE: 2002-06-20
Prior Application removed - See file wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 407
LENGTH: 2609
TYPE: DNA
ORGANISM: Homo Sapien
US-10-176-482-407

Query Match 99.9% Score 1267.4; DB 9; Length 2609;
Best Local Similarity 99.9% Pred. No. 0;
Matches 1268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 CAGTGCAGCAATGGCTTGTGACCTGATGCGCAGTCCAGACAGTGTGATATGATGAA 60
DB 525 CAGTGCAGCAATGGCTTGTGACCTGATGCGCAGTCCAGACAGTGTGATATGATGAA 584
QY 61 TGGCGAACCATCCCGAGGCTGCGGAGAGACATGATGTGTAAACCAAAATGCGGAG 120
DB 585 TGGCGAACCATCCCGAGGCTGCGGAGAGACATGATGTGTAAACCAAAATGCGGAG 644
QY 121 TATTATGCAATTCCTCCGAGAAACCTGTGTATGAGAGGCTTACTGAAACCTTACTG 180
DB 645 TATTATGCAATTCCTCCGAGAAACCTGTGTATGAGAGGCTTACTGAAACCTTACTG 704
QY 181 ACCCTTACTGAGCTGCGTACCGAGAGCTGCGCAGACCTGCTGACCTCAAACTATGCC 240
DB 705 ACCCTTACTGAGCTGCGTACCGAGAGCTGCGCAGACCTGCTGACCTCAAACTATGCC 764

```

US-10-175-738-407
; Sequence 407, Application US/10175738
; Publication No. US2003002294A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zhenlin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C45
; CURRENT APPLICATION NUMBER: US/10/175,738
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; LENGTH: 2609
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-175-738-407

Query Match 99.94; Score 1267.4; DB 9; Length 2609;
Best Local Similarity 99.94; Pred. No. 0;
Matches 1268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAGTCACGAGTGGCTTTGACCTGATGCCAGTCAGAGAGAGTGGTAAATGAGTAA 60
DB 525 CAGTCACGAGTGGCTTTGACCTGATGCCAGTCAGAGAGAGTGGTAAATGAGTAA 584
QY 61 TCCCGACCATCCCGAGGCTCCGAGAGACATGATGTGTAAATGAGGAG 120
DB 585 TCCCGACCATCCCGAGGCTCCGAGAGACATGATGTGTAAATGAGGAG 644
QY 121 TATTTATGATTCCTCCGAGCAACCTGTATCGAGGAGGCTTAACTGAAACCTTACTG 180
DB 645 TATTTATGATTCCTCCGAGCAACCTGTATCGAGGAGGCTTAACTGAAACCTTACTG 704
QY 181 ACCCCCTACTAGGTCCTGATACCCAGAGCTGCCCACTCAGCTCCAACTATGCC 240
DB 705 ACCCCCTACTAGGTCCTGATACCCAGAGCTGCCCACTCAGCTCCAACTATGCC 764
QY 241 ACGATCTCCAGGCTCTTATATGCCCTTTGATACCAATGATGAACCAATGAT 300
DB 765 ACGATCTCCAGGCTCTTATATGTCGCTTTGATACCAATGATGAACCAATGAT 824
QY 301 GTGATGTGAGAGAGTGTCAACAGATTCACCACTGCAACCCAGATGTGATC 360
DB 825 GTGATGTGAGAGAGTGTCAACAGATTCACCACTGCAACCCAGATGTGATC 884
QY 361 AATATGAAAGGGGTACACCTGCTCTGACCCAGAGAGATTTGGCTTGAAGGAGCAG 420
DB 885 AATATGAAAGGGGTACACCTGCTCTGACCCAGAGAGATTTGGCTTGAAGGAGCAG 944
QY 421 TGTATGACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 945 TGTATGACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1004
QY 481 GGATCTCTTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 1005 GGATCTCTTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1064
QY 541 CAATATGAGAGAGTGTGACCAAGAGAACCTCTGCTGCAAACTGCTCAACACCTAC 600
DB 1065 CAATATGAGAGAGTGTGACCAAGAGAACCTCTGCTGCAAACTGCTCAACACCTAC 1124
QY 601 GGCTCTTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660

DB 1125 GGCTCTTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1184
QY 661 AGTATATGAGAGAGTGTGACCAAGAGAACCTCTGCTGCAAACTGCTCAACACCTAC 720
DB 1185 AGTATATGAGAGAGTGTGACCAAGAGAACCTCTGCTGCAAACTGCTCAACACCTAC 1244
QY 721 CCGGACATATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 1245 CCGGACATATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1304
QY 781 TGCAGAGATCAAGAGATGTGAGCAGAGAGACACAGCTGCAACCTGCAAGAGAGTGC 840
DB 1305 TGCAGAGATCAAGAGATGTGAGCAGAGAGACACAGCTGCAACCTGCAAGAGAGTGC 1364
QY 841 TCAATTTCAAG 900
DB 1365 TCAATTTCAAG 1424
QY 901 AGGATCAGTATTAACCGCTGTATGTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB 1425 AGGATCAGTATTAACCGCTGTATGTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1484
QY 961 TTTACCATCTTGTATGAGAGAGATGACAGTGTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB 1485 TTTACCATCTTGTATGAGAGAGATGACAGTGTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1544
QY 1021 TTCAG 1080
DB 1545 TTCAG 1604
QY 1081 GGAATGAG 1140
DB 1605 GGAATGAG 1664
QY 1141 ATGACAG 1200
DB 1665 ATGACAG 1724
QY 1201 AACACTGATCACTTCAAG 1260
DB 1725 AACACTGATCACTTCAAG 1784
QY 1261 TACCATTTC 1269
DB 1785 TACCATTTC 1793

RESULT 11
US-10-175-752-407
; Sequence 407, Application US/10175752
; Publication No. US2003002295A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zhenlin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C60
; CURRENT APPLICATION NUMBER: US/10/175,752
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; LENGTH: 2609

Db 1485 TTACACATCTGTACCGGAGCATGAGACGTGTGTGTCAGACGCTCCGCTCCGCTGACATTC 1544
QY 1021 TTCCAAATGCAAGCAGACACCGGCTACCTGGGGCTATTACATTTTCCAGATCAAAATCT 1080
Db 1545 TTCCAAATGCAAGCAGACACCGGCTACCTGGGGCTATTACATTTTCCAGATCAAAATCT 1604
QY 1081 GGGAAATGAGGCGAGAGATTTTACATGCGGCAAAAGGCGCCCATCATGATCCACCCCTGATG 1140
Db 1605 GGGAAATGAGGCGAGAGATTTTACATGCGGCAAAAGGCGCCCATCATGATCCACCCCTGATG 1664
QY 1141 ATGACAGCCCCCATCAAAAGGCGCGGAAATCCAGCTGAGCTTGGAAATGATCACTGTC 1200
Db 1665 ATGACAGCCCCCATCAAAAGGCGCGGAAATCCAGCTGAGCTTGGAAATGATCACTGTC 1724
QY 1201 AACACGTGATCACTTCAGAGGAGCGCTCCGTGATCCGACCTGGATATATGTGTGCGAG 1260
Db 1725 AACACGTGATCACTTCAGAGGAGCGCTCCGTGATCCGACCTGGATATATGTGTGCGAG 1784
QY 1261 TACCCATTTC 1269
Db 1785 TACCCATTTC 1793

RESULT 9

US-10-173-706-407
; Sequence 407, Application US/10173706
; Publication No. US2003022293A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jilan
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zhenlin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C7
; CURRENT APPLICATION NUMBER: US/10/173,706
; PRIORITY FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 407
; LENGTH: 2609
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-173-706-407

Query Match 99.9% Score 1267.4; DB 9; Length 2609;
Best Local Similarity 99.9% Pred. No. 0;
Matches 1268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAGTGCAGATGCTTTGACCTGGATGCGCAGACAGCACTGTTAGATATTTAGTGA 60
Db 525 CAGTGCAGATGCTTTGACCTGGATGCGCAGACAGCACTGTTAGATATTTAGTGA 584
QY 61 TGGCGAACAATCCCGAGGCTTGGCGAGAGACATGATGTGTAAACCAAAATGCGGG 120
Db 585 TGGCGAACAATCCCGAGGCTTGGCGAGAGACATGATGTGTAAACCAAAATGCGGG 644
QY 121 TATTATGATTCCTCCGAGCAAAACCTGTGTATGAGAGGCGCTTACTGAAACCTTACTG 180
Db 645 TATTATGATTCCTCCGAGCAAAACCTGTGTATGAGAGGCGCTTACTGAAACCTTACTG 704
QY 181 ACCCGCTACGAGTCCGTACCGACAGACAGCTGCGCCACCACTCACTCAAACTTACC 240
Db 705 ACCCGCTACGAGTCCGTACCGACAGACAGCTGCGCCACCACTCACTCAAACTTACC 764
QY 241 ACGATTCACAGGCTCTTATATGCGGCTTGGATACAGATGATGAAGCAACCAATGT 300

Db 765 ACGATTCACAGGCTCTTATATGCGGCTTGGATACAGATGATGAAGCAACCAATGT 824
QY 301 GTGATGTGACAGAGTGTGCAACAGATTCACACAGTGCACCCACCAAGATCTGATC 360
Db 825 GTGATGTGACAGAGTGTGCAACAGATTCACACAGTGCACCCACCAAGATCTGATC 884
QY 361 AATAGTGAAGGCGGTTACACCTGCTCCGACCGAGGATTTGGCTTGTGGAAGCCAG 420
Db 885 AATAGTGAAGGCGGTTACACCTGCTCCGACCGAGGATTTGGCTTGTGGAAGCCAG 944
QY 421 TGTAGACATGATGATGATGCTGATGCTTACTGTCACAGACCTGTCGATGTTCT 480
Db 945 TGTAGACATGATGATGATGCTGATGCTTACTGTCACAGACCTGTCGATGTTCT 1004
QY 481 GGATCTATTCTTGTACATGCAACCCGCTTTTACCCTCAATGAGATGGAAGTCTTG 540
Db 1005 GGATCTATTCTTGTACATGCAACCCGCTTTTACCCTCAATGAGATGGAAGTCTTG 1064
QY 541 CAAGATGGAAGAGTGTGCCACCGAGAACCCCTGCGTCAAACTGGGTCAACACTTAC 600
Db 1065 CAAGATGGAAGAGTGTGCCACCGAGAACCCCTGCGTCAAACTGGGTCAACACTTAC 1124
QY 601 GGCTCTTTTCATCTGCGGCTGTGACCCAGATATGAATGAGAAAGTGGCTTCAATTC 660
Db 1125 GGCTCTTTTCATCTGCGGCTGTGACCCAGATATGAATGAGAAAGTGGCTTCAATTC 1184
QY 661 AGTATATGACAGAGTGTGCTCTGATGCTTCTGTCGCAACATGATGTGTGAACCG 720
Db 1185 AGTATATGACAGAGTGTGCTCTGATGCTTCTGTCGCAACATGATGTGTGAACCG 1244
QY 721 CCGGACATATCTTCTGCTGCTGCGCCTCAGGCTCATCTGCTGATGACCAACCGAG 780
Db 1245 CCGGACATATCTTCTGCTGCTGCGCCTCAGGCTCATCTGCTGATGACCAACCGAG 1304
QY 781 TGCCAGACATCAACGAATGTGACAGACAGAACCAACAGCTGCAACCTGACAGAGCTGC 840
Db 1305 TGCCAGACATCAACGAATGTGACAGACAGAACCAACAGCTGCAACCTGACAGAGCTGC 1364
QY 841 TACAAATTTACAAAGGCGCTTCAAAATGATCAACCCCTGCTGTGAGAGAGCTTATCTG 900
Db 1365 TACAAATTTACAAAGGCGCTTCAAAATGATCAACCCCTGCTGTGAGAGAGCTTATCTG 1424
QY 901 AGATCAGTGTATTAACCGCTGTATGTCTGCTGTAAGAACCTGCTGACAGACAGGCC 960
Db 1425 AGATCAGTGTATTAACCGCTGTATGTCTGCTGTAAGAACCTGCTGACAGACAGGCC 1484
QY 961 TTTACATCTTGTACCGGAGATGAGACGTGTGTGACAGAGCTTCCGCTGACATC 1020
Db 1485 TTTACATCTTGTACCGGAGATGAGACGTGTGTGACAGAGCTTCCGCTGACATC 1544
QY 1021 TTCCAAATGCAAGCAGACACCGGCTGATGCGGCAAAAGGCGCCCATCATGATCCACCTG 1080
Db 1545 TTCCAAATGCAAGCAGACACCGGCTGATGCGGCAAAAGGCGCCCATCATGATCCACCTG 1604
QY 1081 GGGAAATGAGGCGAGAGATTTTACATGCGGCAAAAGGCGCCCATCATGATCCACCTG 1140
Db 1605 GGGAAATGAGGCGAGAGATTTTACATGCGGCAAAAGGCGCCCATCATGATCCACCTG 1664
QY 1141 ATGACAGCCCCCATCAAAAGGCGCGGAAATCCAGCTGAGCTTGGAAATGATCACTGTC 1200
Db 1665 ATGACAGCCCCCATCAAAAGGCGCGGAAATCCAGCTGAGCTTGGAAATGATCACTGTC 1724
QY 1201 AACACTGTCACTTCAGAGGAGCGCTCCGTGATCCGACCTGGATATATGTGTGCGAG 1260
Db 1725 AACACTGTCACTTCAGAGGAGCGCTCCGTGATCCGACCTGGATATATGTGTGCGAG 1784
QY 1261 TACCCATTTC 1269
Db 1785 TACCCATTTC 1793

RESULT 10

DB 765 ACGATCTCCAGGCGCTTATATGCGGCTTTGGATACAGATGATGAAGAACCAACATGT 824
QY 301 GTGGATGTGAGAGTGTGCAAGATTCACACGAGTGAACCCACCCAGATTCGACATC 360
DB 825 GTGATGTGAGAGTGTGCAAGATTCACACGAGTGAACCCACCCAGATTCGACATC 884
QY 361 AATAGTGAAGGCGGCTACCTGCTCTGACACGAGATATGCTTGTGAAGGCGAG 420
DB 885 AATAGTGAAGGCGGCTACCTGCTCTGACACGAGATATGCTTGTGAAGGCGAG 944
QY 421 TGTCTTAACATGATGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 480
DB 945 TGTCTTAACATGATGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1004
QY 481 GGATCTTATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
DB 1005 GGATCTTATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1064
QY 541 CAAGATGTGAACGAGTGTGACACGAGAACCCCTGCTGCAAAACCTGCTCAACACTAC 600
DB 1065 CAAGATGTGAACGAGTGTGACACGAGAACCCCTGCTGCAAAACCTGCTCAACACTAC 1124
QY 601 GGCTCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 1125 GGCTCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1184
QY 661 AGTATGTGAGAGTGTGACACGAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB 1185 AGTATGTGAGAGTGTGACACGAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTG 1244
QY 721 CCGGCGACATACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 1245 CCGGCGACATACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1304
QY 781 TGGCAGACATCAACGATGTGACACGAGAACCCACAGTGTGAACCTGCTGCTGCTG 840
DB 1305 TGGCAGACATCAACGATGTGACACGAGAACCCACAGTGTGAACCTGCTGCTGCTG 1364
QY 841 TACAATTTAAGAGGCGCTTCAATGATGATGATGATGATGATGATGATGATGATG 900
DB 1365 TACAATTTAAGAGGCGCTTCAATGATGATGATGATGATGATGATGATGATGATG 1424
QY 901 AGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
DB 1425 AGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1484
QY 961 TTTACATCTTGTGACGAGGAGTGTGATGATGATGATGATGATGATGATGATGATG 1020
DB 1485 TTTACATCTTGTGACGAGGAGTGTGATGATGATGATGATGATGATGATGATGATG 1544
QY 1021 TTTCAATTTGACAGGCGCTTCAATGATGATGATGATGATGATGATGATGATGATG 1080
DB 1545 TTTCAATTTGACAGGCGCTTCAATGATGATGATGATGATGATGATGATGATGATG 1604
QY 1081 GGGAAATGAGGCGAGAGATTTTACATGCGGCAAAAGGCGCGCTTCAATGATGATG 1140
DB 1605 GGGAAATGAGGCGAGAGATTTTACATGCGGCAAAAGGCGCGCTTCAATGATGATG 1664
QY 1141 ATGACAGGCGCGCTTCAATGAGGCGCGGGAATTCAGTGTGATGATGATGATGATG 1200
DB 1665 ATGACAGGCGCGCTTCAATGAGGCGCGGGAATTCAGTGTGATGATGATGATGATG 1724
QY 1201 AACACTGTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1260
DB 1725 AACACTGTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1784
QY 1261 TACCCATTC 1269
DB 1785 TACCCATTC 1793

RESULT 7
US-10-176-758-407

Sequence 407, Application US/10176758
; Publication No. US2003008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey J.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C104
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 407
; LENGTH: 2609
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-758-407

Query Match 99.9%; Score 1267.4; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAGTGCAGATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 60
DB 525 CAGTGCAGATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 584
QY 61 TGGCAGACATCAACGATGTGACACGAGAACCCACAGTGTGAACCTGCTGCTGCTG 120
DB 585 TGGCAGACATCAACGATGTGACACGAGAACCCACAGTGTGAACCTGCTGCTGCTG 644
QY 121 TATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
DB 645 TATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 704
QY 181 ACCCTTACTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
DB 705 ACCCTTACTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 764
QY 241 ACGATGTGAGGCGCTTATATGCGGCTTTGGATACAGATGATGATGATGATGATG 300
DB 765 ACGATGTGAGGCGCTTATATGCGGCTTTGGATACAGATGATGATGATGATGATG 824
QY 301 GTGATGTGAGGAGTGTGCAACAGATTTCCACAGTGTGCAACCCACAGATGATGATG 360
DB 825 GTGATGTGAGGAGTGTGCAACAGATTTCCACAGTGTGCAACCCACAGATGATGATG 884
QY 361 AATAGTGAAGGCGGCTACCTGCTCTGACACGAGATATGCTTGTGAAGGCGAG 420
DB 885 AATAGTGAAGGCGGCTACCTGCTCTGACACGAGATATGCTTGTGAAGGCGAG 944
QY 421 TGTCTTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
DB 945 TGTCTTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1004
QY 481 GGATCTTATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
DB 1005 GGATCTTATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1064
QY 541 CAAGATGTGAACGAGTGTGACACGAGAACCCCTGCTGCAAAACCTGCTCAACACTAC 600
DB 1065 CAAGATGTGAACGAGTGTGACACGAGAACCCCTGCTGCAAAACCTGCTCAACACTAC 1124
QY 601 GGCTCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660

PRIOR APPLICATION NUMBER: PCT/US99/21547

Query Match 99.9%; Score 1267.4; DB 9; Length 2609;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 CAGTGCACGAATGGCTTTGACCTGGATCGCCAGTACAGAGACAGTGTATGATTTGATGAA 60
 525 CAGTGCACGAATGGCTTTGACCTGGATCGCCAGTACAGAGAGTGTATGATTTGATGAA 584
 61 TGGCCAGACCATCCCGAGGCGCTGCCAGAGACATGATGTGTGTTAACCAAAATGGCGGG 120
 585 TGGCCAGACCATCCCGAGGCGCTGCCAGAGACATGATGTGTGTTAACCAAAATGGCGGG 644
 121 TATTTATGCAATCCCGGAGCAAAACCTGTATGAGAGGGCCCTACTGCAACCCCTACTCG 180
 645 TATTTATGCAATCCCGGAGCAAAACCTGTATGAGAGGGCCCTACTGCAACCCCTACTCG 704
 181 ACCCCCTACTGAGTCCGTACCCAGCAGCTGCCACACACTCTGAGCTCCAACTATGCC 240
 705 ACCCCCTACTGAGTCCGTACCCAGCAGCTGCCACACACTCTGAGCTCCAACTATGCC 764
 241 AGCATCTCCAGGCGCTTTATATGCGCGCTTGGATACCATGATGATGAAGCAACCAATGT 300
 765 AGCATCTCCAGGCGCTTTATATGCGCGCTTGGATACCATGATGATGAAGCAACCAATGT 824
 301 GTGATGTGGAGAGTGTGCAACAGATTCACACAGTGCACAAACCCACCCCAATCTGCATC 360
 825 GTGATGTGGAGAGTGTGCAACAGATTCACACAGTGCACAAACCCACCCCAATCTGCATC 884
 361 AATACTGAAGGGGGGTACACCTGCTCTGCAACGACGAGATTTGGCTTCTGGAAGGCCAG 420
 885 AATACTGAAGGGGGGTACACCTGCTCTGCAACGACGAGATTTGGCTTCTGGAAGGCCAG 944
 421 TGGCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
 945 TGGCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1004
 481 GGATCCATATCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
 1005 GGATCCATATCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1064
 541 CAAGATGTGAACGAGTGTGCCACGAGAACCCCTGCGTGCACAACTGCTGCAACACCTAC 600
 1065 CAAGATGTGAACGAGTGTGCCACGAGAACCCCTGCGTGCACAACTGCTGCAACACCTAC 1124
 601 GGCTCTTTATATGCTGCGCGCTGTGACCCAGATATGATGATGATGATGATGATGATGATG 660
 1125 GGCTCTTTATATGCTGCGCGCTGTGACCCAGATATGATGATGATGATGATGATGATGATG 1184
 661 AGTATATGAGAGAGTGTGCCACGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 720
 1185 AGTATATGAGAGAGTGTGCCACGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1244
 721 CCCGGACATATCTTGTGCTGCTGCGCTGACCCAGGCTACATCTGCTGATGATGATGATGATG 780
 1245 CCCGGACATATCTTGTGCTGCTGCGCTGACCCAGGCTACATCTGCTGATGATGATGATGATG 1304
 781 TGGCAAGACATCAAGATGTGACGACAGAACCAACAGTGCACCTGCAACGACAGCTGC 840
 1305 TGGCAAGACATCAAGATGTGACGACAGAACCAACAGTGCACCTGCAACGACAGCTGC 1364
 841 TACAAATTTACAAGGGGGCTTCAAAATGATGATGATGATGATGATGATGATGATGATGATG 900
 1365 TACAAATTTACAAGGGGGCTTCAAAATGATGATGATGATGATGATGATGATGATGATGATG 1424
 901 AGGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
 1425 AGGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1484
 961 TTTTACATCTTGTACCGGAGACATGACGATGATGATGATGATGATGATGATGATGATGATG 1020
 1485 TTTTACATCTTGTACCGGAGACATGACGATGATGATGATGATGATGATGATGATGATGATG 1544

1021 TTCCAAATGCAACCCAGACCCGCTACCTTGAGGCGCTATTACATTTTCCAGATCAAAATCT 1080
 1545 TTCCAAATGCAACCCAGACCCGCTACCTTGAGGCGCTATTACATTTTCCAGATCAAAATCT 1604
 1081 GGGATGAGGGGAGAGATTTTACATGCGGCAAAAGGGGCCCATCAGTGCACCCCTGGTG 1140
 1605 GGGATGAGGGGAGAGATTTTACATGCGGCAAAAGGGGCCCATCAGTGCACCCCTGGTG 1664
 1141 ATGACAGGCCCATCAAAAGGGGCCCGGAAATCCAGCTGACCTTGAATGATGATGATGATG 1200
 1665 ATGACAGGCCCATCAAAAGGGGCCCGGAAATCCAGCTGACCTTGAATGATGATGATGATG 1724
 1201 AACACTGATCAACTTCAAGAGCAGTCCGCTGATCCGACTGCGGATATATGTGTGCGAG 1260
 1725 AACACTGATCAACTTCAAGAGCAGTCCGCTGATCCGACTGCGGATATATGTGTGCGAG 1784
 1261 TACCCATTC 1269
 1785 TACCCATTC 1793

RESULT 6 US-10-174-590-407

Sequence 407, Application US/10174590
 Publication No. US20030008352A1

GENERAL INFORMATION:
 APPLICANT: Baker, Kevin P.
 APPLICANT: Chen, Jian
 APPLICANT: Desnoyers, Luc
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Pan, James
 APPLICANT: Smith, Victoria
 APPLICANT: Watanabe, Colin K.
 APPLICANT: Wood, William I.
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 TITLE OF INVENTION: ACIDS ENCODING THE SAME
 FILE REFERENCE: P3430R1C42
 CURRENT APPLICATION NUMBER: US/10/174,590
 CURRENT FILING DATE: 2002-06-18
 Prior application removed - See File Wrapper or Palm
 NUMBER OF SEQ ID NOS: 612
 SEQ ID NO 407
 LENGTH: 2609
 TYPE: DNA
 ORGANISM: Homo Sapien
 US-10-174-590-407

Query Match 99.9%; Score 1267.4; DB 9; Length 2609;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 CAGTGCACGAATGGCTTTGACCTGGATCGCCAGTACAGAGACAGTGTATGATTTGATGAA 60
 525 CAGTGCACGAATGGCTTTGACCTGGATCGCCAGTACAGAGAGTGTATGATTTGATGAA 584
 61 TGGCCAGACCATCCCGAGGCGCTGCCAGAGACATGATGTGTAAACCAAAATGGCGGG 120
 585 TGGCCAGACCATCCCGAGGCGCTGCCAGAGACATGATGTGTAAACCAAAATGGCGGG 644
 121 TATTTATGCAATCCCGGAGCAAAACCTGTATGAGAGGGCCCTACTGCAACCCCTACTCG 180
 645 TATTTATGCAATCCCGGAGCAAAACCTGTATGAGAGGGCCCTACTGCAACCCCTACTCG 704
 181 ACCCCCTACTGAGTCCGTACCCAGCAGCTGCCACACACTCTGACGCTCCAACTATGCC 240
 705 ACCCCCTACTGAGTCCGTACCCAGCAGCTGCCACACACTCTGACGCTCCAACTATGCC 764
 241 AGCATCTCCAGGCGCTTTATATGCGCGCTTGGATACCATGATGATGAAGCAACCAATGT 300

PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066840
PRIOR FILING DATE: 1997-11-25
PRIOR APPLICATION NUMBER: 60/069694
PRIOR FILING DATE: 1997-12-16
PRIOR APPLICATION NUMBER: 60/074086
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/074092
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/095998
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/097000
PRIOR FILING DATE: 1998-08-18
PRIOR APPLICATION NUMBER: 60/099601
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099811
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100858
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101922
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/106032
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: 60/109304
PRIOR FILING DATE: 1998-11-20
PRIOR APPLICATION NUMBER: 60/125778
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/139695
PRIOR FILING DATE: 1999-06-15
PRIOR APPLICATION NUMBER: 60/145070
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/149396
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/169495
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 08/918874
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 08/933821
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 08/960507
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 08/114844
PRIOR FILING DATE: 1998-07-14
PRIOR APPLICATION NUMBER: 09/136801
PRIOR FILING DATE: 1998-08-19
PRIOR APPLICATION NUMBER: 09/136804
PRIOR FILING DATE: 1998-08-19
PRIOR APPLICATION NUMBER: 09/136828
PRIOR FILING DATE: 1998-08-19
PRIOR APPLICATION NUMBER: 09/158342
PRIOR FILING DATE: 1998-09-21
PRIOR APPLICATION NUMBER: 09/180997
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 09/202088
PRIOR FILING DATE: 1998-12-08
PRIOR APPLICATION NUMBER: 09/254311
PRIOR FILING DATE: 1999-03-03
PRIOR APPLICATION NUMBER: 09/254460
PRIOR FILING DATE: 1999-03-09
PRIOR APPLICATION NUMBER: 09/254465
PRIOR FILING DATE: 1999-03-05
PRIOR APPLICATION NUMBER: 09/284663

PRIOR FILING DATE: 1999-04-15
PRIOR APPLICATION NUMBER: 09/332928
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/332929
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/333075
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/333077
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380138
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380139
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/403296
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/403297
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/423741
PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: 09/423844
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 09/522342
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: 09/548815
PRIOR FILING DATE: 2000-04-13
PRIOR APPLICATION NUMBER: 09/664610
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: 09/665350
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: 09/709238
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 09/767609
PRIOR FILING DATE: 2001-01-22
PRIOR APPLICATION NUMBER: 09/802706
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 09/808689
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 09/866028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 09/870574
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: 09/872035
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 09/886342
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: PCT/US98/14552
PRIOR FILING DATE: 1998-07-14
PRIOR APPLICATION NUMBER: PCT/US98/18824
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: PCT/US98/19093
PRIOR FILING DATE: 1998-09-14
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: PCT/US98/19437
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: PCT/US98/24855
PRIOR FILING DATE: 1998-11-20
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: 1998-12-01
PRIOR APPLICATION NUMBER: PCT/US98/25190
PRIOR FILING DATE: 1998-11-25
PRIOR APPLICATION NUMBER: PCT/US99/05028
PRIOR FILING DATE: 1999-03-08
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/20111
PRIOR FILING DATE: 1999-09-01
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15

Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CAGTGCAGATGGCTTTGACCTGGATGCCAGTACAGACAGTGTATGATGAA 60
Db 457 CAGTGCAGATGGCTTTGACCTGGATGCCAGTACAGACAGTGTATGATGAA 516
QY 61 TGCCGAACCATCCCGAGAGCTGCGAGAGACATGATGTGTAAACCAAAATGGCGG 120
Db 517 TGCGAACAACATCCCGAGAGCTGCGAGAGACATGATGTGTAAACCAAAATGGCGG 576
QY 121 TATTTATGATTCCTCCGAGCAAAACCTGTGTATGAGAGGCCCTACTCGAACCCCTACTCG 180
Db 577 TATTTATGATTCCTCCGAGCAAAACCTGTGTATGAGAGGCCCTACTCGAACCCCTACTCG 636
QY 181 ACCCCCTACTGAGTCCGATACCGAGAGCTCCCGACACTCTGAGCTCCAAACTATCC 240
Db 637 ACCCCCTACTGAGTCCGATACCGAGAGCTCCCGACACTCTGAGCTCCAAACTATCC 696
QY 241 AGCATCTCAGAGCCCTTATATGCGCGTTTGATACAGATGATGAAGCAACCATGT 300
Db 697 AGCATCTCAGAGCCCTTATATGCGCGTTTGATACAGATGATGAAGCAACCATGT 756
QY 301 GTGATGTGACAGAGTGTGCAACAGATTTCCACAGTGCACACCCACAGATCTGATC 360
Db 757 GTGATGTGACAGAGTGTGCAACAGATTTCCACAGTGCACACCCACAGATCTGATC 816
QY 361 AATACTGAAGCGGCTACCTGCTGCGACCGAGAGATTTGCGCTCTGGAAGCCAG 420
Db 817 AATACTGAAGCGGCTACCTGCTGCGACCGAGATTTGCGCTCTGGAAGCCAG 876
QY 421 TCGTTAGACATTTGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 877 TCGTTAGACATTTGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 936
QY 481 GGATCTATCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Db 937 GGATCTATCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 996
QY 541 CAAGATGTGACAGAGTGTGCAACAGATTTCCACAGTGCACACCCACAGATCTGATC 600
Db 997 CAAGATGTGACAGAGTGTGCAACAGATTTCCACAGTGCACACCCACAGATCTGATC 1056
QY 601 GGCTCTTTATCTGCGCGCTGTGACCCAGATGATGATGATGATGATGATGATGATGAT 660
Db 1057 GGCTCTTTATCTGCGCGCTGTGACCCAGATGATGATGATGATGATGATGATGATGAT 1116
QY 661 AGTATGATGACAGAGTGTGCAACAGATTTCCACAGTGCACACCCACAGATCTGATC 720
Db 1117 AGTATGATGACAGAGTGTGCAACAGATTTCCACAGTGCACACCCACAGATCTGATC 1176
QY 721 CCCGCGACATCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 1177 CCCGCGACATCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1236
QY 781 TGCCGAAGATCAACCAATGTGAGACAGACACAGATGATGATGATGATGATGATGATGAT 840
Db 1237 TGCCGAAGATCAACCAATGTGAGACAGACACAGATGATGATGATGATGATGATGAT 1296
QY 841 TACAAATTTACAAGGGGCTTCAAAATGATGATGATGATGATGATGATGATGATGATGAT 900
Db 1297 TACAAATTTACAAGGGGCTTCAAAATGATGATGATGATGATGATGATGATGATGATGAT 1356
QY 901 AGGATCAGTATTAACCGCTGTATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 1357 AGGATCAGTATTAACCGCTGTATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1416
QY 961 TTTACCATCTTGTACCGGAGCATGACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db 1417 TTTACCATCTTGTACCGGAGCATGACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1476
QY 1021 TTCCAAATGCAAGCAGACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 1477 TTCCAAATGCAAGCAGACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1536
```

```
QY 1081 GGGATGAGGAGGAGAGATTTTACATGCGGCAAAAGGCGCCCATCATGATGCCACCTGGTG 1140
Db 1537 GGGATGAGGAGGAGAGATTTTACATGCGGCAAAAGGCGCCCATCATGATGCCACCTGGTG 1596
QY 1141 ATGACAGCGCCCATCAAGAGGCGCCCGGAAATCCAGCTGACCTTGGAAATATCATCTGTC 1200
Db 1597 ATGACAGCGCCCATCAAGAGGCGCCCGGAAATCCAGCTGACCTTGGAAATATCATCTGTC 1656
QY 1201 AACACTGTCACTCACTGACAGAGCAGCTCCGATCCGACCTGGGATATGATGTCGAG 1260
Db 1657 AACACTGTCACTCACTGACAGAGCAGCTCCGATCCGACCTGGGATATGATGTCGAG 1716
QY 1261 TACCATTC 1269
Db 1717 TACCATTC 1725

RESULT 5
US-10-066-500-14
; Sequence 14, Application US/10066500
; Patent No. US20020177165A1
; GENERAL INFORMATION:
; APPLICANT: Avi J. Ashkenazi
; APPLICANT: Kevin P. Baker
; APPLICANT: David A. Bolstein
; APPLICANT: Luc Desnoyers
; APPLICANT: Dan L. Baton
; APPLICANT: Napoleone Ferrara
; APPLICANT: Sherman Fong
; APPLICANT: Wei-Qiang Gao
; APPLICANT: Hanspeter Gerber
; APPLICANT: Mary E. Gerritsen
; APPLICANT: Audrey Goddard
; APPLICANT: Paul J. Godowski
; APPLICANT: Austin L. Guiney
; APPLICANT: Ivar J. Kljavin
; APPLICANT: Jennie P. Mather
; APPLICANT: Mary A. Napier
; APPLICANT: James Pan
; APPLICANT: Nicholas F. Paoni
; APPLICANT: Margaret Ann Roy
; APPLICANT: Timothy A. Stewart
; APPLICANT: Daniel Tumas
; APPLICANT: Colin K. Watanabe
; APPLICANT: P. Mickey Williams
; APPLICANT: William I. Wood
; APPLICANT: Zemin Zang
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3130R1C7
; CURRENT APPLICATION NUMBER: US/10/066, 500
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 10/002,796
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062285
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063733
; PRIOR FILING DATE: 1997-10-29
```

US-09-083-002-1

Query Match 100.0% Score 1269; DB 10; Length 2362;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGTGCAGAAATGCTTTGACCTGATGCCAGTGCAGAGAGTGTATAGTATGATGAA 60
 DB CAGTGCAGAAATGCTTTGACCTGATGCCAGTGCAGAGAGTGTATAGTATGATGAA 320
 QY 61 TGGCGAAGCAATCCCGAGAGCCCTGCGAGAGAGACATGATGTGTATACCAAAATGCGGG 120
 DB TGGCGAAGCAATCCCGAGAGCCCTGCGAGAGAGACATGATGTGTATACCAAAATGCGGG 380
 QY 121 TATTTATGATCTCCCGGACAAACCTGTGTATGAGGGCCCTACTCGAAACCCCTACTCG 180
 DB TATTTATGATCTCCCGGACAAACCTGTGTATGAGGGCCCTACTCGAAACCCCTACTCG 440
 QY 181 ACCCCCTACTAGTCCGTACCCAGAGCTGCCCAACCACTGTCAGCTCCAACTATCCC 240
 DB ACCCCCTACTAGTCCGTACCCAGAGCTGCCCAACCACTGTCAGCTCCAACTATCCC 500
 QY 241 ACGATCTCCAGGCGCTTATATGCGCGTTTGGATACCAATGATGAAACCAATGT 300
 DB ACGATCTCCAGGCGCTTATATGCGCGTTTGGATACCAATGATGAAACCAATGT 560
 QY 301 GTGATGTGAGAGAGTGTGCAACAGATTCACCAAGTGAACCCCAAGATCTGATC 360
 DB GTGATGTGAGAGAGTGTGCAACAGATTCACCAAGTGAACCCCAAGATCTGATC 620
 QY 361 AATACTGAAGCGGGTACACCTGCTCTGACACGAGAGATATGAGCTTTGGAAGCGAG 420
 DB AATACTGAAGCGGGTACACCTGCTCTGACACGAGAGATATGAGCTTTGGAAGCGAG 680
 QY 421 TGCCTAAGCATTTGATGATGTGCTATGTTACTGTCGACAGCTCTGTGCAATTTCT 480
 DB TGCCTAAGCATTTGATGATGTGCTATGTTACTGTCGACAGCTCTGTGCAATTTCT 740
 QY 481 GGATCTATTTCTTGTACATGCAACCTGGTTTACCCCAATGAGATGGAAGTCTTGC 540
 DB GGATCTATTTCTTGTACATGCAACCTGGTTTACCCCAATGAGATGGAAGTCTTGC 800
 QY 541 CAAGATGTGAAGAGTGTGCAACGAGAACCCCTGGTGCACCAACCTGCGTCAACACTTAC 600
 DB CAAGATGTGAAGAGTGTGCAACGAGAACCCCTGGTGCACCAACCTGCGTCAACACTTAC 860
 QY 601 GGCCTCTTACCTGCGCGTGTGACCCAGATATGAACCTTGAGAGAGATGCGCTTCAATGC 660
 DB GGCCTCTTACCTGCGCGTGTGACCCAGATATGAACCTTGAGAGAGATGCGCTTCAATGC 920
 QY 661 AGTATATGAGAGAGTGTGACCTGCTGTAGTCCCTGCAACATGAGTGTGAACCAAG 720
 DB AGTATATGAGAGAGTGTGACCTGCTGTAGTCCCTGCAACATGAGTGTGAACCAAG 980
 QY 721 CCCGCAATATCTTGTCTCTGCGCCCTCAGGCTACATCTCTGTGATGACCAACGAGC 780
 DB CCCGCAATATCTTGTCTCTGCGCCCTCAGGCTACATCTCTGTGATGACCAACGAGC 1040
 QY 781 TGGCAAGACATCAAGATGTGAGACAGAGAACACACCTGCAACCTGCAAGAGAGCTGC 840
 DB TGGCAAGACATCAAGATGTGAGACAGAGAACACACCTGCAACCTGCAAGAGAGCTGC 1100
 QY 841 TACAATTTTACAAGGGGGCTTCAATATGATGACACCCATCCGCTGTGAGAGAGCTTATCTG 900
 DB TACAATTTTACAAGGGGGCTTCAATATGATGACACCCATCCGCTGTGAGAGAGCTTATCTG 1160
 QY 901 AGGATCAAGTATTAACCGCTGTATGTCTGCTGAGAACCTTGGCTGACAGAGACAGCC 960
 DB AGGATCAAGTATTAACCGCTGTATGTCTGCTGAGAACCTTGGCTGACAGAGACAGCC 1220
 QY 961 TTTTACCATCTTGTACCGGAGACATGACGTGTGTGAGAGACCTTCCCGTGTGACATC 1020
 DB TTTTACCATCTTGTGTACCGGAGACATGACGTGTGTGAGAGACCTTCCCGTGTGACATC 1280

QY 1021 TTCCAAATGCAAGCCAGCAGCCGCTACCCCTGAGGCGCTATTACATTTCCAGATCAATCT 1080
 DB TTCCAAATGCAAGCCAGCAGCAGCCGCTACCCCTGAGGCGCTATTACATTTCCAGATCAATCT 1340
 QY 1081 GGGATGAGAGGCGAGAGAAATTTTACATGCGGCAAAAGCGGCCCAATCAGTGCACCTGTG 1140
 DB GGGATGAGAGGCGAGAGAAATTTTACATGCGGCAAAAGCGGCCCAATCAGTGCACCTGTG 1400
 QY 1141 ATGACAGCCCGCATTAAGAGCGCCCGGAAATCCAGCTGAGACTTGGAAATGATCAGCTGC 1200
 DB ATGACAGCCCGCATTAAGAGCGCGCCGAAATCCAGCTGAGACTTGGAAATGATCAGCTGC 1460
 QY 1201 AACACTGTATCAACTTCAGAGGAGAGCTCCGTATGCGAGTGGATATATGTGTGCGAG 1260
 DB AACACTGTATCAACTTCAGAGGAGAGCTCCGTATGCGAGTGGATATATGTGTGCGAG 1520
 QY 1261 TACCCATTC 1269
 DB TACCCATTC 1529

RESULT 4

US-09-836-561-2
 Sequence 2, Application US/09836561
 Patent No. US20020038006A1
 GENERAL INFORMATION:
 APPLICANT: Bandman, Olga
 Corley, Neil C.
 Guejler, Karl J.
 TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/836,561
 FILING DATE: 16-Apr-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/212,168
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0333 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 TELEX: <Unknown>
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2550 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: CORNNOT01
 CLONE: 45517
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-09-836-561-2

Query Match 100.0% Score 1269; DB 10; Length 2550;
 Best Local Similarity 100.0%; Pred. No. 0;

TOPOLOGY: linear
MOLECULE TYPE: CDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-041-016-1

Query Match 100.0%; Score 1269; DB 9; Length 2362;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CAGTACAGATGCTTTGACCTGGATGCCAGTCAGACAGTGTATATATATATGAA 60
DB 261 CAGTACAGATGCTTTGACCTGGATGCCAGTCAGACAGTGTATATATATATGAA 320
QY 61 TGCCGAACATCCCGGAGGCGCTGGAGAGACATGATGTGTATACCAAAATGCGGG 120
DB 321 TGCCGAACATCCCGGAGGCGCTGGAGAGACATGATGTGTATACCAAAATGCGGG 380
QY 121 TATTTATGCAATTCCTCCGAGCAAAACCTGTGTATCGAGGGCCCTACTGAAACCCCTACTCG 180
DB 381 TATTTATGCAATTCCTCCGAGCAAAACCTGTGTATCGAGGGCCCTACTGAAACCCCTACTCG 440
QY 181 ACCCCCTACTGAGTCCGTACCCAGACAGTGTGCCCCACCACTCTCAGCTCAAACTATCCC 240
DB 441 ACCCCCTACTGAGTCCGTACCCAGACAGTGTGCCCCACCACTCTCAGCTCAAACTATCCC 500
QY 241 ACGATCTCCAGGCGCTTTATATAGCCGCTTTGGATACCAAGATGATGAAGCAACCAATGT 300
DB 501 ACGATCTCCAGGCGCTTTATATAGCCGCTTTGGATACCAAGATGATGAAGCAACCAATGT 560
QY 301 GTGATGTGGACAGAGTGTGCAACAGATTTCCACACAGTGCACACCCAGCCAGATCTGCATC 360
DB 561 GTGATGTGGACAGAGTGTGCAACAGATTTCCACACAGTGCACACCCAGATCTGCATC 620
QY 361 AATCTGAAGGCGGCTACACCTGCTCTGCAACCGAGATATGTGCTTCTGGAAGCGCAG 420
DB 621 AATCTGAAGGCGGCTACACCTGCTCTGCAACCGAGATATGTGCTTCTGGAAGCGCAG 680
QY 421 TGCTTAGACATTTGATGATGTGCTATGCTACTGCGACAGCTGTGCGAATGTTCTT 480
DB 681 TGCTTAGACATTTGATGATGTGCTATGCTACTGCGACAGCTGTGCGAATGTTCTT 740
QY 481 GGAATCCATTTCTGATGATGCAACCCGCTGTTTACCTCCATATGAGATGGAAGGTCTTGC 540
DB 741 GGAATCCATTTCTGATGATGCAACCCGCTGTTTACCTCCATATGAGATGGAAGGTCTTGC 800
QY 541 CAAGATGTGAACAGAGTGTGCCACCGAGAAACCCCTGCGTCAAACTGCGTCAACACCTAC 600
DB 801 CAAGATGTGAACAGAGTGTGCCACCGAGAAACCCCTGCGTCAAACTGCGTCAACACCTAC 860
QY 601 GGCTCTTTCTATCTGCGCTGTGACCCAGAGATATGAACCTTGAAGAGATGGCGTTCAATGC 660
DB 861 GGCTCTTTCTATCTGCGCTGTGACCCAGAGATATGAACCTTGAAGAGATGGCGTTCAATGC 920
QY 661 AGTATATGAGAGAGTGTGACCTTCTGAGTCTCTGCGCAACATGATGTGTAACAG 720
DB 921 AGTATATGAGAGAGTGTGACCTTCTGAGTCTCTGCGCAACATGATGTGTAACAG 980
QY 721 CCCGGACATATCTTGTGCTCTGCGCTCAAGGCTACATCTCTGTGTGATGACACCGAAGC 780
DB 981 CCCGGACATATCTTGTGCTCTGCGCTCAAGGCTACATCTCTGTGTGATGACACCGAAGC 1040
QY 781 TGCCAGACATCAACCAATGTGAGCACAGAACACAGTGCACCTGACAGACGTCG 840
DB 1041 TGCCAGACATCAACCAATGTGAGCACAGAACACAGTGCACCTGACAGACGTCG 1100
QY 841 TACAATTTACAAAGGGGCTTCAATGATGACACCCATCGCTGTGAGAGCCCTATCTG 900
DB 1101 TACAATTTACAAAGGGGCTTCAATGATGACACCCATCGCTGTGAGAGCCCTATCTG 1160
QY 901 AGGATCAGTATTAACCGCTGTATGTCTGCTGAGAACCCCTGCTCAGAGACACAGCC 960
DB 1161 AGGATCAGTATTAACCGCTGTATGTCTGCTGAGAACCCCTGCTCAGAGACACAGCC 1220

```

```

QY 961 TTACCATCTTGTACCGGACATGAGACGCTGTGTGACGACGCTCCGTTCCCGCTGACATC 1020
DB 1221 TTACCATCTTGTACCGGACATGAGACGCTGTGTGACGACGCTCCGTTCCCGCTGACATC 1280
QY 1021 TTCCAAATGCAAGCAGACACCGCTTACCTGTGGGCTTATACATTTTCCAGATCAAAATCT 1080
DB 1281 TTCCAAATGCAAGCAGACACCGCTTACCTGTGGGCTTATACATTTTCCAGATCAAAATCT 1340
QY 1081 GGGATGAGGCGACAGATTTTACATGCGGCAACGGGCCCCATAGTGCACCTGGTG 1140
DB 1341 GGGATGAGGCGACAGATTTTACATGCGGCAACGGGCCCCATAGTGCACCTGGTG 1400
QY 1141 ATGACACGCCCATCAAAAGGCGCCGGGAAATCCAGCTGACCTTGAATGATCACTGTC 1200
DB 1401 ATGACACGCCCATCAAAAGGCGCCGGGAAATCCAGCTGACCTTGAATGATCACTGTC 1460
QY 1201 AACACTGTCATCAACTTCAGAGGAGCTCCCTGTATCCGACCTGGGATATATGTCTGCGAG 1260
DB 1461 AACACTGTCATCAACTTCAGAGGAGCTCCCTGTATCCGACCTGGGATATATGTCTGCGAG 1520
QY 1261 TACCATTTC 1269
DB 1521 TACCATTTC 1529

```

RESULT 3
US-09-083-002-1
Sequence 1, Application US/09083002
Patent No. US20010016650A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Racie, Lisa A.
APPLICANT: Lavallee, Edward R.
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Evans, Cheryl
APPLICANT: Agostino, Michael
APPLICANT: Lu, ZhiJian
APPLICANT: Honjo, Tasuku
APPLICANT: Tashiro, Kei
APPLICANT: Nakamura, Tomoyuki
TITLE OF INVENTION: SECRETED PROTEINS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U. S. A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083,002
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: P-41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-3851
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2362 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA

MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 211..1554
 US-09-275-805-1

Query Match 100.0%; Score 1269; DB 10; Length 1717;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGTGCAGAAATGCTTTGACCTGGATGCCAGTACAGACAGTGTATATTTGATGA 60
 DB 286 CAGTGCAGAAATGCTTTGACCTGGATGCCAGTACAGACAGTGTATATTTGATGA 345
 QY 61 TGGCGAACAATCCCGAGAGCTGGCGAGAGACATGATGTGTAAACCAATGGCGGG 120
 DB 346 TGGCGAACAATCCCGAGAGCTGGCGAGAGACATGATGTGTAAACCAATGGCGGG 405
 QY 121 TATTTATGCAATCCCGAGACAAACCTGTGTATGAGGGCCCTACTGAAACCCCTACTCG 180
 DB 406 TATTTATGCAATCCCGAGACAAACCTGTGTATGAGGGCCCTACTGAAACCCCTACTCG 465
 QY 181 ACCCCCTACACAGTCCGCTACCCAGACAGTCCGCTACCCAGACAGTCCGCTACCCAGAC 240
 DB 466 ACCCCCTACACAGTCCGCTACCCAGACAGTCCGCTACCCAGACAGTCCGCTACCCAGAC 525
 QY 241 ACATCTCCAGAGCTCTTATATGCGCGCTTTGGATACAGATGATGAAGAACCAATGT 300
 DB 526 ACATCTCCAGAGCTCTTATATGCGCGCTTTGGATACAGATGATGAAGAACCAATGT 585
 QY 301 GTGGAGTGGACGAGTGTGCAACAGATTCACACAGTGCACACCCAGACAGTGTGCATC 360
 DB 586 GTGGAGTGGACGAGTGTGCAACAGATTCACACAGTGCACACCCAGACAGTGTGCATC 645
 QY 361 AATACGTAAGGGGGGTACACTGCTCCGCAACCGAGATATGCTTGTGAAGCCAG 420
 DB 646 AATACGTAAGGGGGGTACACTGCTCCGCAACCGAGATATGCTTGTGAAGCCAG 705
 QY 421 TGGTTACATTTGATGATGATGCTGATGATGATGATGATGATGATGATGATGATGATG 480
 DB 706 TGGTTACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 765
 QY 481 GGATCTTATTTCTGTACATGCAACCCCTGTGTTTACCTCAATGAGATGGAAGTCTTGC 540
 DB 766 GGATCTTATTTCTGTACATGCAACCCCTGTGTTTACCTCAATGAGATGGAAGTCTTGC 825
 QY 541 CAAGATGTAAGCGAGTGTGCAACCGAGAACCCCTGCGTGAACCTGCTCAACACCTAC 600
 DB 826 CAAGATGTAAGCGAGTGTGCAACCGAGAACCCCTGCGTGAACCTGCTCAACACCTAC 885
 QY 601 GGCCTCTTCAATCTCCGCTGTGACCCAGATATGAACTTGAAGAGATGGCGTTCTTTCG 660
 DB 886 GGCCTCTTCAATCTCCGCTGTGACCCAGATATGAACTTGAAGAGATGGCGTTCTTTCG 945
 QY 661 AGTGATATGACGAGTGTGCAAGTGTCTCTGCTCCCAACATGAGTGTGAACGAG 720
 DB 946 AGTGATATGACGAGTGTGCAAGTGTCTCTGCTCCCAACATGAGTGTGAACGAG 1005
 QY 721 CCGGCGACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 DB 1006 CCGGCGACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1065
 QY 781 TGGCAAGACATCAAGATGATGACACAGAACACAGTGTGCAACCTGAGAGAGAGAGAG 840
 DB 1066 TGGCAAGACATCAAGATGATGACACAGAACACAGTGTGCAACCTGAGAGAGAGAGAG 1125
 QY 841 TACAATTTACAAGGGGCTTCAAAATGATGCAACCCCATCCGCTGTGAGAGAGCTTATCTG 900
 DB 1126 TACAATTTACAAGGGGCTTCAAAATGATGCAACCCCATCCGCTGTGAGAGAGCTTATCTG 1185
 QY 901 AGGATAGTATTAACCCCTGTATGTCTGCTGGAACCCCTGCTGCAAGAGACAGCC 960
 DB 1186 AGGATAGTATTAACCCCTGTATGTCTGCTGGAACCCCTGCTGCAAGAGACAGCC 1245

QY 961 TTTCACATCTGTACCCGGAGACATGAGCGGTGTGACAGACGCTCCGCTCCGCTGACATC 1020
 DB 1246 TTTCACATCTGTACCCGGAGACATGAGCGGTGTGACAGACGCTCCGCTCCGCTGACATC 1305
 QY 1021 TTCCAAATGCAAGCCAGACACCGCTACCCCTGGGGCTTATATTTTCCAGATCAAAATCT 1080
 DB 1306 TTCCAAATGCAAGCCAGACACCGCTACCCCTGGGGCTTATATTTTCCAGATCAAAATCT 1365
 QY 1081 GGGATATGAGGCGACAGAAATTTTACATGCGGGCAAAAGGGGCCCATCATAGTCCACCCCTGGTG 1140
 DB 1366 GGGATATGAGGCGACAGAAATTTTACATGCGGGCAAAAGGGGCCCATCATAGTCCACCCCTGGTG 1425
 QY 1141 ATGACAGCCCATCAAGAGGCGCCCGGAAATCCAGCTGATGATGGAATGATCACTGTC 1200
 DB 1426 ATGACAGCCCATCAAGAGGCGCCCGGAAATCCAGCTGATGATGGAATGATCACTGTC 1485
 QY 1201 AACACTGTATCAACTTCAGAGGACGTCCTGTATCCGACTCGGATATATGTGTGCGAG 1260
 DB 1486 AACACTGTATCAACTTCAGAGGACGTCCTGTATCCGACTCGGATATATGTGTGCGAG 1545
 QY 1261 TACCCATTG 1269
 DB 1546 TACCCATTG 1554

RESULT 2
 US-10-041-016-1
 Sequence 1, Application US/10041016
 Patent No. US20020165151A1
 GENERAL INFORMATION:
 APPLICANT: Jacobs, Kenneth
 McCoy, John M.
 Racie, Lisa A.
 Lavalie, Edward R.
 Werberg, David
 Treacy, Maurice
 Evans, Cheryl
 Agostino, Michael
 Lu, Zhijian
 Honjo, Tasuku
 TITLE OF INVENTION: SECRETED PROTEINS
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESSES:
 ADDRESS: Genetics Institute, Inc.
 STREET: 87 Cambridgepark Drive
 CITY: Cambridge
 STATE: MA
 COUNTRY: U.S.A.
 ZIP: 02140
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/041, 016
 FILING DATE: 07-Jan-2002
 CLASSIFICATION: <unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/083, 002
 FILING DATE: 21-MAR-1998
 ATTORNEY/AGENT INFORMATION:
 NAME: Sprunger, Suzanne A.
 REGISTRATION NUMBER: P-41,323
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 498-8284
 TELEFAX: (617) 876-5851
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2362 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 11:41:03 ; Search time 206.749 Seconds

9581.191 million cell updates/sec

File: 14-US-09-674-379A-15
Perfect score: 1269
Sequence: 1 cagtgcacgaatgcttga.....atgtctgcagtraccattc 1265

scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1085931 seqs, 780495707 residues

Total number of hits satisfying chosen parameters: 2171862

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match	0%
------------------	---------------	----

Listing first 45 summaries

```

1: /cgn2_6/p/odataa/1/pub/pna/us07_PUBCOMB.seq.*
2: /cgn2_6/p/odataa/1/pub/pna/PCTI_NEW_PUB.seq.*
3: /cgn2_6/p/odataa/1/pub/pna/US06_NEW_PUB.seq.*
4: /cgn2_6/p/odataa/1/pub/pna/US06_PUBCOMB.seq.*
5: /cgn2_6/p/odataa/1/pub/pna/US07_NEW_PUB.seq.*
6: /cgn2_6/p/odataa/1/pub/pna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/p/odataa/1/pub/pna/US08_NEW_PUB.seq.*
8: /cgn2_6/p/odataa/1/pub/pna/US08_PUBCOMB.seq.*
9: /cgn2_6/p/odataa/1/pub/pna/US09_NEW_PUB.seq.*
10: /cgn2_6/p/odataa/1/pub/pna/US09_PUBCOMB.seq.*
11: /cgn2_6/p/odataa/1/pub/pna/US10_NEW_PUB.seq.*
12: /cgn2_6/p/odataa/1/pub/pna/US10_PUBCOMB.seq.*
13: /cgn2_6/p/odataa/1/pub/pna/US60_NEW_PUB.seq.*
14: /cgn2_6/p/odataa/1/pub/pna/US60_PUBCOMB.seq.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1269	100.0	1717	10	US-09-275-805-1	Sequence 1, App1
2	1269	100.0	2362	9	US-10-041-016-1	Sequence 1, App1
3	1269	100.0	2362	10	US-09-083-002-1	Sequence 1, App1
4	1269	100.0	2550	10	US-09-836-561-2	Sequence 2, App1
5	1267.4	99.9	2609	9	US-10-066-500-14	Sequence 14, App1
6	1267.4	99.9	2609	9	US-10-174-590-407	Sequence 407, App
7	1267.4	99.9	2609	9	US-10-176-738-407	Sequence 407, App
8	1267.4	99.9	2609	9	US-10-175-737-407	Sequence 407, App
9	1267.4	99.9	2609	9	US-10-173-706-407	Sequence 407, App
10	1267.4	99.9	2609	9	US-10-175-738-407	Sequence 407, App
11	1267.4	99.9	2609	9	US-10-175-752-407	Sequence 407, App
12	1267.4	99.9	2609	9	US-10-176-483-407	Sequence 407, App
13	1267.4	99.9	2609	9	US-10-176-757-407	Sequence 407, App
14	1267.4	99.9	2609	9	US-10-176-913-407	Sequence 407, App
15	1267.4	99.9	2609	9	US-10-180-552-407	Sequence 407, App
16	1267.4	99.9	2609	9	US-10-180-557-407	Sequence 407, App
17	1267.4	99.9	2609	9	US-10-173-700-407	Sequence 407, App
18	1267.4	99.9	2609	9	US-10-174-572-407	Sequence 407, App
19	1267.4	99.9	2609	9	US-10-174-579-407	Sequence 407, App

20	1267.4	99.9	2609	9	US-10-174-588-407	Sequence 407, App
21	1267.4	99.9	2609	9	US-10-174-588-407	Sequence 407, App
22	1267.4	99.9	2609	9	US-10-175-733-407	Sequence 407, App
23	1267.4	99.9	2609	9	US-10-175-740-407	Sequence 407, App
24	1267.4	99.9	2609	9	US-10-175-743-407	Sequence 407, App
25	1267.4	99.9	2609	9	US-10-176-488-407	Sequence 407, App
26	1267.4	99.9	2609	9	US-10-176-492-407	Sequence 407, App
27	1267.4	99.9	2609	9	US-10-176-747-407	Sequence 407, App
28	1267.4	99.9	2609	9	US-10-176-750-407	Sequence 407, App
29	1267.4	99.9	2609	9	US-10-176-985-407	Sequence 407, App
30	1267.4	99.9	2609	9	US-10-176-991-407	Sequence 407, App
31	1267.4	99.9	2609	9	US-10-176-991-407	Sequence 407, App
32	1267.4	99.9	2609	9	US-10-176-992-407	Sequence 407, App
33	1267.4	99.9	2609	9	US-10-176-993-407	Sequence 407, App
34	1267.4	99.9	2609	9	US-10-184-658-407	Sequence 407, App
35	1267.4	99.9	2609	9	US-10-002-796-14	Sequence 14, App
36	1267.4	99.9	2609	9	US-10-066-473-14	Sequence 14, App
37	1267.4	99.9	2609	9	US-10-106-494-14	Sequence 14, App
38	1267.4	99.9	2609	9	US-10-173-691-407	Sequence 407, App
39	1267.4	99.9	2609	9	US-10-173-692-407	Sequence 407, App
40	1267.4	99.9	2609	9	US-10-173-705-407	Sequence 407, App
41	1267.4	99.9	2609	9	US-10-174-576-407	Sequence 407, App
42	1267.4	99.9	2609	9	US-10-174-585-407	Sequence 407, App
43	1267.4	99.9	2609	9	US-10-174-586-407	Sequence 407, App
44	1267.4	99.9	2609	9	US-10-175-747-407	Sequence 407, App
45	1267.4	99.9	2609	9	US-10-176-481-407	Sequence 407, App

ALIGNMENTS

```

RESULT 1
US-09-275-805-1
: Sequence 1, Application US/09275805
: Patent No. US20010051358A1
: GENERAL INFORMATION:
: APPLICANT: OLSEN, HENRIK S.
: APPLICANT: LI, HAODONG
: TITLE OF INVENTION: EXTRACELLULAR EPIDERMAL GROWTH FACTOR
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: HUMAN GENOME SCIENCES, INC.
: STREET: 9410 KEY WEST AVENUE
: CITY: ROCKVILLE
: STATE: MD
: COUNTRY: US
: ZIP: 20850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/275,805
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/839,525
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: BROOKES, ANDERS A.
: REGISTRATION NUMBER: 36,373
: REFERENCE/DOCKET NUMBER: PF224
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 309-8509
: TELEFAX: (301) 309-8512
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1717 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear

```


Sat Jul 5 15:31:33 2003

us-09-674-379a-15.rni

Page 11

```

;      FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..3753
;
PCT-US95-02251-2

```

Query Match	4.18;	Score 51.6;	DB 5;	Length 3753;
Best Local Similarity	49.18;	Pred. No. 6.8e-06;		

Matches 166; Conservative 0; Mismatches 169; Indels 3; Gaps 1.

Qy 466 TGTGGCAATGTTCCCTGGATCCTATTCTTGTATCATGCAACCCCTGGTTTACCCCTCAATGAG 525
Db 2212 TGTGGAAAACTTCCCGGGTTCTTACCGTTGCACGTGTGGCCAGGGGATACGAACCCGACA 2271
Qy 526 GATGNAAGGCTCTGGCAAGANTGNAACGAGTGTGCCACCGAAMCCCTGC --GTGCA 582
Db 2272 GGACCCCTCAGTTTGCATATGACCTGGATACCTGTGGGCTGGGAAAGTGTGCAGAGATGGC 2331
Qy 583 ACCTGCGTCAACACCTTACGGCTCTTTCATCTGCCGCTGACCCAGATATGAACCTTGAG 642
Db 2332 ATCTGCAGAACACACAGCGCTCTTCCAGTGTAGTGCCTCTCCGGCTATCAATCTGTCA 2391
Qy 643 GAAGATGGCGTTCAATGTCAGTATATGAGACAGTGCAGCTCTGTGATTCCTGTGCCA 702
Db 2392 AGGAGTCCGAGCCGCTGTGAGGACATGATGATGAATGTACTTCCCTGGCGCCTGCATGGG 2451
Qy 703 CATGAGTGTGGAACCAAGCCCGGCAATACTTCTGCTCTGCCCTCCAGGCTCAATCTG 762
Db 2452 GGTACATGCAATTAACCAATGTGTTCTCAAGATGTCTGTGCCCTGGGTCAATCGGTTG 2511
Qy 763 CTGATGACAACCGAGCTGCACAGATCAACAGATG 800
Db 2512 GTGGGCGGACGAGAGTGCAGAAAGATATGATGAGTG 2549

Search completed: July 3, 2003, 17:40:02
Job time : 59.5574 secs

QY 835 ACCTGCTACATTTTACAAGGGGGCTTCAATGATGACCCATCCGCTGTGAGAGCCT 894
DB 3481 GAATGCAAGAACACAGAAAGGTTCTTACCAATGCTCTGTACACAGGCGTTCCAGCTGTC 3540
QY 895 TATCTGAGATGATGATTAACCCGCTGTATGTCTCTGCTGAGAACCCCTGG 944
DB 3541 AATGGCACCATGTGTGAGAGCGTGAATGATGTGTGGGGAAGACCATTTG 3590

RESULT 14

PCT-US95-02251-17
Sequence 17, Application PC/TUS9502251
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING BONE
TITLE OF INVENTION: CELLS
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: Patentln Release #1.0, Version
SOFTWARE: #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02251
FILING DATE: CONCURRENTLY HERewith
CLASSIFICATION:
PRIOR APPLICATION NUMBER: US 08/316,650
FILING DATE: 30-SEP-1994
CLASSIFICATION:
APPLICATION NUMBER: US 08/199,780
FILING DATE: 18-FEB-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UMIC009P--
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 5502 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "DNA"
FEATURE:
NAME/KEY: CDS
LOCATION: 1..5502
PCT-US95-02251-17

Query Match 4.1%; Score 52.4; DB 5; Length 5502;
Best Local Similarity 47.4%; Pred. No. 4.8e-06;
Matches 223; Conservative 0; Mismatches 241; Indels 6; Gaps 2;

QY 481 GGATCTATTCTTGTATCATGCAACCGCTTTTACCTCAATGAGAGTGAAGTCTTGC 540
DB 3121 GGCTCCCTTTGATGCTCTGTGAGCCGGCTATGAGTACCCAGACAAAGAGGCTGC 3180
QY 541 CAAGATGTGAACGATGCGACGACGAGAACCCCTG---CGTCAANAACGGCTCAACAC 597
DB 3181 CGAGATGTGACGATGTGCGACGACGCTGTGCGCCACGCGGCTCTGCTCAACACG 3240

QY 598 TAGGCTCTTTCATCTGCG---CTGTGACCCAGATATGAATTGAGGAAGATGGCCTT 654
DB 3241 GAGGCTCTTTCATCTGCG---CTGTGACCCAGATATGAATTGAGGAAGATGGCCTT 3300
QY 655 CATTGAGATGATGAGAGTGTGAGTCTTCTGAGTTCCTTGGCAACATGATGTGTG 714
DB 3301 GCCTGTGAAGACTTGTGATGATGTGCTTCCCTGTGAGTGTGCTTCCACAGGCTGTGACC 3360
QY 715 AACCAAGCCGCGACATCTTGTGCTTCCCTGTGAGTGTGCTTCCACAGGCTGTGACC 774
DB 3361 AATAGTGTGAGCTCTTCTGCTGAGAGCTGTGACGAGGCTTACCGCCCAACCCCTG 3420
QY 775 CGAGCTGCCAATCAACGAATGTGAGACAGACAGAACACAGCTGCAACCTGACAGC 834
DB 3421 GGCACAGATGCGAAGATGTGATGATGATGATGATGATGATGATGATGATGATGATG 3480
QY 835 ACCTGCTACATTTTACAAGGGGGCTTCAATGATGACCCATCCGCTGTGAGAGCCT 894
DB 3481 GAATGCAAGAACACAGAAAGGTTCTTACCAATGCTCTGTGACACAGGCGTTCCAGCTGTC 3540
QY 895 TATCTGAGATGATGATTAACCCGCTGTATGTCTCTGCTGAGAACCCCTGG 944
DB 3541 AATGGCACCATGTGTGAGAGCGTGAATGATGTGTGGGGAAGACCATTTG 3590

RESULT 15

PCT-US95-02251-2
Sequence 2, Application PC/TUS9502251
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING BONE
TITLE OF INVENTION: CELLS
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: Patentln Release #1.0, Version
SOFTWARE: #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02251
FILING DATE: CONCURRENTLY HERewith
CLASSIFICATION:
PRIOR APPLICATION NUMBER: US 08/316,650
FILING DATE: 30-SEP-1994
CLASSIFICATION:
APPLICATION NUMBER: US 08/199,780
FILING DATE: 18-FEB-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UMIC009P--
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3753 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,722B
FILING DATE: 07-JUN-1995
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US PCT/US95/02251
FILING DATE: 21-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/316,650
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/199,780
FILING DATE: 18-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: 4100.000500/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (713) 934-7000
TELEFAX: (713) 934-7011
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3759 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3759
US-08-479-722B-3

Query Match 4.2%; Score 53.2; DB 3; Length 3759;
Best Local Similarity 49.4%; Pred. No. 2.1e-06;
Matches 167; Conservative 0; Mismatches 168; Indels 3; Gaps 1;
QY 466 TGTGGAGTCTCCGATGATCTTCTTGTACATGCAACCCGTGTTTACCTCAATGAG 525
DB 2215 TGTGAGAACCTTCGGGCTTCTTACCGTTCACGTGCGCCGAGGATGCAACCCGACA 2274
QY 536 GATGGAAGCTCTTGGCAGATGTGAAGAGTGTGCCAGCAACCCCTGC---GTGCAA 582
DB 2275 GGACGGCTCAGTTGATACACGATGATGATGAGGTGGGAAAGTGTCCCAAGATGCC 2334
QY 583 ACCTGCGTCAACACTAGCGCTCTTTCATCTGCCGCTGTGACCCAGATATGAATTGAG 642
DB 2335 ATCTGACACGACACACAGGCTCTTTCAGATGTCAAGTGTCCGCTATCATCTGTCA 2394
QY 643 GAAATGGGCTTCATGAGTGTATGAGAGTGTGACAGTCTTCTGTAGTTCCTCTGCCA 702
DB 2395 AGGAGTCGAGCGCGCTGTGAGACATGATGATGATGATGATGATGATGATGATGATG 2454
QY 703 CATGATGTGTAAACAGCCCGGACATATCTTCTGCTCCGCTCCAGGCTCATGCTG 762
DB 2455 GGTGACTGATCATATACCAATGTTCTCTACAGATGTCTGTGCTCCCTGGGTCTAGG 2514
QY 763 CTGATGACACCGAAGCTGCCAAGATCAACGAATG 800
DB 2515 GTGGGCGGCGAAGTGAAGAAAGATATGATGAGTG 2552

RESULT 13
US-08-479-722B-1
Sequence 1, Application US/08479722B
Patent No. 6074840
GENERAL INFORMATION:
APPLICANT: Bonadio, Jeffrey
APPLICANT: Yin, Wushan
TITLE OF INVENTION: LATENT TGF β BINDING PROTEIN (LTBP)
NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:
ADDRESSEE: Williams, Morgan & Amerson
STREET: 7676 Hillmont, Suite 250
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77040
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,722B
FILING DATE: 07-JUN-1995
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US PCT/US95/02251
FILING DATE: 21-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/316,650
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/199,780
FILING DATE: 18-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: 4100.000500/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (713) 934-7000
TELEFAX: (713) 934-7011
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5499 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: /desc = "DNA"
DESCRIPTION:
FEATURE:
NAME/KEY: CDS
LOCATION: 1..5499
US-08-479-722B-1

Query Match 4.1%; Score 52.4; DB 3; Length 5499;
Best Local Similarity 47.4%; Pred. No. 4.8e-06;
Matches 223; Conservative 0; Mismatches 241; Indels 6; Gaps 2;
QY 481 GGATCTATCTTGTACATGCAACCCGTGTTTACCCCAATGAGATGGAAGTCTTGC 540
DB 3121 GGCTCTTATAGATGCTCTGTGAGCGGGCTATGAGTGCACCCAGACAAAGAGGCTGC 3180
QY 541 CAGATGTGAAGAGTGTGCCAGCAGAGACCCCTG---CGTCAAACTGCTCAACACC 597
DB 3181 CGAGATGTGAAGAGTGTGCCAGCAGAGACCCCTGCTGCTCCAGCGGCTGCTCAACAG 3240
QY 598 TAGGCTCTTTCATCTGCGG---CTGTACCCAGATATGAATTGAGAGATGCGGTT 654
DB 3241 GAGGGCTCTCTACCTGCTCAGCGCTGTGAGAGCGGCTACTGGGTGAAGATGCACT 3300
QY 655 CATGCACTGATGAGAGAGTGTGAGAGTCTTCTGTAGTCTCTGTGCAACATGAGTGTG 714
DB 3301 GCCTGTGAAGACTTGATGAAATGTGCTTCCCTGTGAGTGTGCCCAAGAGGCTGTGACC 3360
QY 715 AACCAAGCCGACATATCTGCTCTGCGCTCCAGGCTACATCTGCTGATGATCAAC 774
DB 3361 AATACTGTAGGCTCTCTTCTGCTGCAAGAGATGTACACAGGAGGCTACCGCCCAACCCCTG 3420
QY 775 CGAAGCTGCAAGACATCAACGAATGTGAGCAGCAGAAACACACGCTGCAACCTGACAG 834
DB 3421 GCGAAGAGATCGAAGATGTGATGATGATGATGATGATGATGATGATGATGATGATG 834

```

Db      213 GCCATGAGACATCTCCAGATACAGGCCACACATATTATGCCAACACCATCATACTTT 272
QY      1068 CCAGATCAATCTGGAGATGAGGGCAGAGAAATTTTACATGGCGCAAGGCCCATGAG 1127
Db      273 TCGGATTAATCTGGAAATGAAATGAGAGTCTACTTACGACAAACCCCTGTAG 331
QY      1128 TGGCACCCTGGTGTATGACAGCCCATCAAAAGGGCCCGGAAATCCAGCTGAGTGA 1187
Db      332 TGCATGCTGTGCTGTGTAAGNCATTTATGAGACCAAGAAACATATCTGAGACTGGA 391
QY      1188 AATGATCAC 1196
Db      392 GATGCTGAC 400

```

```

RESULT 10
US-09-404-879A-340/C
; Sequence 340, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 340
; LENGTH: 220
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(220)
; OTHER INFORMATION: n = A,T,C or G
US-09-404-879A-340

```

```

Query Match      6.3%; Score 80.2; DB 4; Length 220;
Best Local Similarity 60.5%; Pred. No. 1.4e-15;
Matches 150; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

```

```

QY      298 TGTGTGATGTGAGCGATGTGCAACAGATTCACCAGCCGCAACCCACAGATTCG 357
Db      219 TGTGTGATGTGAGCGATGTGCAACAGATTCACCAGCCGCAACCCACAGATTCG 160
QY      358 ATCAAACTGAGAGCGGGGTACACCTGCTCTGACACCGAGGATATTGGCTTCTGAAAGC 417
Db      159 CATAACTGCTGCTCTCTATCATCTGACACCTGCTGATGGTTACCGCAAGATCGGGCCC 100
QY      418 CAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 477
Db      99 GAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 40
QY      478 CCGTGAATCTTATCTTATCATCTGATGATGATGATGATGATGATGATGATGATGAT 512
Db      39 CCGTGAATCTTATCTTATCATCTGATGATGATGATGATGATGATGATGATGATGAT 5

```

```

RESULT 11
US-08-282-141-1
; Sequence 1, Application US/08282141
; Patent No. 5538861
; GENERAL INFORMATION:
; APPLICANT: Schneider, Claudio
; APPLICANT: Varnum, Brian
; APPLICANT: Avanzini, Giancarlo
; APPLICANT: Brancolini, Claudio
; APPLICANT: Manfioletti, Guisalberto
; TITLE OF INVENTION: Stimulating Factor for the AXI Receptor
; NUMBER OF SEQUENCES: 4

```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehavenland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: United States
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/282,141
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2461 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
US-08-282-141-1

```

```

Query Match      4.6%; Score 58.2; DB 1; Length 2461;
Best Local Similarity 54.5%; Pred. No. 4.5e-08;
Matches 164; Conservative 0; Mismatches 128; Indels 9; Gaps 2;

```

```

QY      446 ATGCTTACTGCCAGACCTCTGTGCGAATGTTCTCTGATTCCTATTCTGTATCATCAAC 505
Db      625 ACGGGGCTGCTCTCAAGATCTGCCCAACAAGCGGGGTAGCTTCCACATCTCTCCAC 684
QY      506 CTGCTTACCCTCAATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 565
Db      685 GCGGCTTGAGCTCTCTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 744
QY      566 AGAAGCCCTGCTGCAAACTGCTCAACACTA---CGGCTCTTCACTGCGCCCTGTG 622
Db      745 CGAGAGCTGCTGCGGAGAGCGGCTGCAAGAACCTGCGGCTCTGCTCTGCTG 804
QY      623 ACCAGGATATGAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 682
Db      805 ACGAGGCTTGTGCTTACAGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 859
QY      683 TCTGTGATCTCTCTGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 742
Db      860 -TGTGAGGCGCTGTGAGCAGGTCTGCTGAATCCCCAGGAGCTACACCTGCACCT 918
QY      743 G 743
Db      919 G 919

```

```

RESULT 12
US-08-479-722B-3
; Sequence 3, Application US/08479722B
; Patent No. 6074840
; GENERAL INFORMATION:
; APPLICANT: Bonadio, Jeffrey
; APPLICANT: Yin, Wushan
; TITLE OF INVENTION: LATENT TGF( BINDING PROTEIN (LTBP)
; TITLE OF INVENTION: GENES, COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Williams, Morgan & Amerson
; STREET: 7676 Hillmont, Suite 250
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77040
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

QY	1108	CGGGAAGAGGGCCCCCATGAGGCCACCTGGTGTATGACACGCCCCCATTAAGAGGGCCCCGG	1167
Db	1237	CGACAAACAAATCCTGTTAAGTGCATAGCTTGTTCGTGAAGTATTTTCAGGACCAAGA	1296
QY	1168	GAATTCACAGCTGGAGCTTGGAAATGATCATCTGCTCAACACTGTCATTCACCTTAGAGGCAGC	1227
Db	1297	GAACTATTCGTGGACCTGGAGATCTGTCACTCAGCAGTAAAGGACCTTCGCCACAAGC	1356
QY	1328	TCCCTGATCCGACTGCGGATATATGTGTCGCGAGTACCATT	1268
Db	1357	TCTGTGTTAAGATTGACATATAATAGTGGGCCATTTCATT	1397

RESULT 8

```

US-08-897-443-2
Sequence 2, Application US/08897443
Patent No. 5981263
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
APPLICANT: Shah, Purvi
APPLICANT: Kaser, Mathew
TITLE OF INVENTION: HUMAN MATRILIN-3
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/897,443
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0348 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3373 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: UTRSN0102
CLONE: 681719
US-08-897-443-2

```

Query Match	8.4%	Score 106.8;	DB 2;	Length 3373;
Best Local Similarity	57.7%	Pred. No. 3..e-23;		
Matches 211; Conservative	0;	Mismatches 152;	Indels 3;	Gaps 1;

OY	452	ACGCGCAGCACTCTGTGGCCAAATGTTCTTGATCTTATTCTTGTAACAATACCCTGGT	511
Ddb	920	ACTGTGAGAGGCTCTGTGTGAATGTGCCGGGCTCTTCGTCTGCACAGTCTAACATGGCT	979
OY	512	TTACCCCTCAATGAGATGAGAAGTCTTGGCAAGATGTGAACGAAGTGTGCCACCGAAGC	571
Ddb	980	ACGCCCTGGCTAGAGATGGGAGAAAGGCTGTGCTGTGGCACTACTGTGGCTTCACAAAAAC	1039
OY			

QY	572	CC---TGGGTCAAAACCGGTGCAACACCTTACGGCTCTTATCTGACCGCGGTGACCCAG	628
Db	1040	ACGGATGTGACACATGAGTGTCTAAATGCTGTATGGCTCTTACTCTTGCCAGTCCATGAA	1095
QY	629	GATATGAACCTTGAGGAAGATGGCGCTTACCTATTCGACGTATATGAGAGAGTGCAGCTTTCGTG	688
Db	1100	GATTTGGCTCTTAAACCCAGATAAAAAAAACGTGCACAAAGATAGACTACTGTGCTCATATTA	1159
QY	689	AGTTCTCTGTGCAACATGATGTGTGTGAACGACCGCGGCACATACTCTTGTGCTCGGCCCTC	748
Db	1160	ATCACGGATGTGCACACAGAGTGTGTAAACACAGATGATTCCTATTTCGTGCCACATGCCGTA	1219
QY	749	CAGGCTACATCTCTCTGGATGACCAACCGAAGCTCCCAAGACATCAAGAAATGTGAGACA	808
Db	1220	AAGGCTTTAACCTGATTCAGATTAAGAAACCCTGCAGAAGGATCACTACTGTGCATCTGA	1279
QY	809	GGAAAC	814
Db	1280	ACAAAC	1285

RESULT 9

```

US-09-643-597-278
Sequence 278, Application US/09643597
Patent No. 6426072
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Fan, Liqun
APPLICANT: Kalos, Michael D.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Hosken, Nancy
APPLICANT: Fanger, Gary R.
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Henderson, Robert A.
APPLICANT: McNeill, Patricia D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
OF LUNG CANCER
FILE REFERENCE: 210121.455C11
CURRENT APPLICATION NUMBER: US/09/643,597
CURRENT FILING DATE: 2000-08-21
NUMBER OF SEQ ID NOS: 369
SOFTWARE: FastSeq for Windows Version 3.0.
SEQ ID NO 278
LENGTH: 401
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(401)
OTHER INFORMATION: n = A,T,C or G
US-09-643-597-278

```

Query Match	7.0%;	Score 89.4;	DB 4;	Length 401;
Best Local Similarity	54.2%;	Pred. No. 2.6e-18;		
Matches 200; Conservative	0;	Mismatches 168;	Indels 1;	Gaps 1

[illegible]

SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 259
 LENGTH: 1018
 TYPE: DNA
 ORGANISM: Human
 US-09-188-930-259

Query Match
 Best Local Similarity 62.3%; Pred. No. 2.4e-89;
 Matches 515; Conservative 0; Mismatches 309; Indels 3; Gaps 1;

24.9%; Score 316.6; DB 3; Length 1018;
 443 GCTATGTTACTCCAGCAGCTCTGTGCGAATGTTCTGATTCCTATTCCTGATGCA 502
 1 GCTACCGCTACTGCGACACCGCTGCGTGAACCTGCGTGGCTCTCCGCGCCAGTGG 60
 503 ACCCTGTTTACCCTCAATGAGATGGAAGCTTGTGCAAGATGTGAACAGTGTCCA 562
 61 AGCCGGGCTTCCAGCTGGGGCTTACAAACCGCTCTGTGTGATGTGAACGATGTGCA 120
 563 CCGAGAACCCCTGCTCAACCTGCTCAACACCTACGCTTTCATCTGCTGGTGG 622
 121 TGGGGGCCCCATGCGAGCAGCGCTGCTTCAACTCTATGAGACCTTCTGTGCTGCC 180
 623 ACCCAGATATGAACTTGAAGAGATGAGCGTTCAATTCAGATGATGAGCAGTCACT 682
 181 ACCAGGCTATGAGCTCATGCGGATGGCTTCTCTGAGATGATGATGATGATGATG 240
 683 TCTCTAGTCTCTGCTGCAACATGATGTGTGAACCGCCGCGGCACTATCTCTCTCT 742
 241 ACTCTAGCTACTCTGTGCTATACCGCTGCTCAACGAGCAGCGCTTCTCTCTCACT 300
 743 GCCCTCAGGCTACATCTCTGTGATGACAAACGAGCTGCCAAGCATCAAGATGTG 802
 301 GCCACAGGGTATACAGCTGTGTG--GCCACACGCTTGGCAAGACATGATGATG 357
 803 AGCAGAGAACACACAGCTGCAACCTGCGACGAGCTGTGATTAATTAAGGGGCTTCA 862
 358 AGCTGTGTGCGACACAGCTGTGCGAGGCGCCAAACCTGTGTCAACTTCAATGGGGCTTAC 417
 863 AATGATCAGACCCCTGCTGTGAGAGCCTTATCTGAGATGATGATGATGATGATGATG 922
 418 GCTGCTGAGACACCAACCGCTGCTGAGAGCCCTTACATCCAGTCTCTGAGAACCGCTGTG 477
 923 TGTGCTCTGTGAGAACCTGTGCTGAGAGACGAGCCCTTACATCTGTGATACGGGACA 982
 478 TGTGCGCGGCTCCAACTCTATGATGTGAGAGAGCCCTTACATCTGTGAGAACCGCTTAC 537
 983 TGAAGTGTGTCTGAGACGCTCCGCTTCCGCTGACATCTTCCAAATGCAAGCCAGACCC 1042
 538 TGAACATACACCTGAGAGCGGAGCGTGCCTGCTGACGTGTCCAGATCCAGAGCGACCTCCG 597
 1043 GCTACCCCTGGGGCTATTACATTTTCCAGATCAATTCAGATGAGGAGGAGAGATTTT 1102
 598 TCTACCCCGCTGCTTCAATGCTCTTCAATGCTGCTGAGAACTCGAGGGGGGACTTTT 657
 1103 ACATGGGCAAAAGGGCCCATCAGTGCACACCTGCTGTGATGACAGCCCATCAAAAGGCG 1162
 658 ACATATGAGCAAAATCAACAGCTGAGCGCATGTGCTCTGCGCGCGGCTGAGAGGGGCG 717
 1163 CCGGGGAATTCAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1222
 718 CCGGGGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 777
 1223 GAGAGTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1269
 778 CCAAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 824

RESULT 7
 US-09-248-757-1
 Sequence 1, Application us/09248757
 Patent No. 6417342
 GENERAL INFORMATION:

APPLICANT: STONE, EDWIN M.
 APPLICANT: SHEFFIELD, VAL C.
 TITLE OF INVENTION: MACULAR DEGENERATION DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: UIA-018.02
 CURRENT APPLICATION NUMBER: US/09/248.757
 CURRENT FILING DATE: 1999-02-12
 NUMBER OF SEQ ID NOS: 4
 SOFTWARE: Patent In Ver. 2.1
 SEQ ID NO 1
 LENGTH: 2512
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-248-757-1

Query Match
 Best Local Similarity 19.0%; Score 241; DB 4; Length 2512;
 Matches 547; Conservative 0; Mismatches 445; Indels 9; Gaps 3;

271 GATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 330
 403 GGTACAGCAAGTGAACACACAGTGTGCCAAGACATAGAGAGTGTGACAGGAGC 462
 331 CACCAAGTCAACCCACAGATCTGATCAATATCTGAAGCGGGTACACCTGTCTGC 390
 463 CACCAAGTGAAGCAGACCAAGTGTGATCAATTTTACCGGGATCTTGTGATGATGATG 522
 391 ACCGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 447
 523 CTCTCTGATATCAAGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 582
 448 GGTACTGCGAGCAGCTCTGTGCGAATGTTCTGTGATGATGATGATGATGATGATG 507
 583 CATATATGCGACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 642
 508 GGTATTCCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 567
 643 GGTATTCATTTGGCAGAACCACTATACCTGCTGATGATGATGATGATGATGATGATG 702
 568 AACCCCTGCTGCAACCTGCTGCAACCTGCTGCAACCTGCTGCAACCTGCTGCAACCT 627
 703 AATCAATGCTGCTGCAACCTGCTGCAACCTGCTGCAACCTGCTGCAACCTGCTGCAAC 762
 628 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 687
 763 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 822
 688 GATATTCCTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 747
 823 AGCTACTGCTGCAATATCAATATGATGATGATGATGATGATGATGATGATGATGATG 882
 748 CCAAGCTATACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 807
 883 CAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 939
 808 AGGAGACCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 867
 940 ACAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 996
 868 ATGAGACCCATCCGCTGTGAGAGCCTTATCTGAGATGATGATGATGATGATGATGATG 927
 997 TATTCAGGAATCTTGTCAAGATCCCTTACATTTTCAACCAAGAACCAAGCAAGGATG 1056
 928 CTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 987
 1057 CCAATGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1116
 988 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1047
 1117 ATCCGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1176
 1048 CTTGGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1107
 1177 GCCAACCAATCAATATCTTCTGATGATGATGATGATGATGATGATGATGATGATGATG 1236

Db	487	GCCTCCAGGCTWCTTCTTTGGAGATTAACGGAAGCTGCCAGATATCAATGAATGTG	546
QY	803	AGCAGAGAACCCACAGTGTCAACCTGTGACAGACAGCTGTCTACAATTTACAAAGGGGGCTTCA	862
Db	547	AGCMCCGGAACACACATGACTACTCCCCGTGCACACTTGTCAATCTGCAAGGGGGCTTCA	606
QY	863	AATGATGAGACCCCATCCGGCTGTGAGGAGCCCTTATCTGAGATCAGTGAATACCGCTGT	922
Db	607	AATGATGAGACCCCATCTGTGCGAGAGCCCTTATCTGTGATGTGGGATTAACCGCTGT	666
QY	923	TGTGTCTGTGTGAGAACCTGTGGCTGTACAGAGACACCCCTTTACATCTTGTACCCGGACA	982
Db	667	TGTGCCCTGTGTGAGAACTGTGGCTGTACAGAGACACCATTAACATCTTGTTCGGGACA	726
QY	983	TGGAACGTGTGTACAGAGAGCTTCCTCCCGCTGACATCTTCCAATGTCAAGCAGCAGCC	104
Db	727	TGGATGTGTGTACAGAGAGCTTCCTCCCGCTGACATCTTCCAATGTCAAGCAACACACC	786
QY	1043	GCTACCCCTGGGGCCCTATTACATTTTCCAGATCAATGTGGGAATGGAGGACAGAAATTT	110
Db	787	GATACCCCTGGGGCCCTATTACATTTTCCAGATCAATGTGGGAATGGAGGCTGAGATCT	846
QY	1103	ACATGCGGCAAAAGGGGCCCATCAGTGCACCCCTGTGTGATGACAGGCCCATCAAGGGC	116
Db	847	ACATGCGGCAAAAGGGGCCCTTACATGTGCACCCCTGTGTGATGACAGGCCCATCAAGGGC	906
QY	1163	CCCGGGAATTCAGCTGTGATTTGGAATGATCAGTGCAACTGTGATCAATCACTTCAGAG	122
Db	907	CTCGGAGATTCAGCTGTGATTTGGAATGATCAGTGCAACTGTGATCAATCACTTCAGAG	966
QY	1223	GCAGCTCCGTGATCCGACTGGCGATATATGTGTGCGACATCCCAATTC	1269
Db	967	GCAGCTCCGTGATCCGACTGGCGATATATGTGTGCGACATCCCAATTC	1013

RESULT 4
US-08-833-963C-1
Sequence 1, Application US/08833963C
Patent No. 5916769
GENERAL INFORMATION:
APPLICANT: Olsén, et al.
TITLE OF INVENTION: Extracellular/Epidermal Growth Factor
TITLE OF INVENTION: HCBAB58X
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Ave
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833, 963C
FILING DATE: 11-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US96/05033
FILING DATE: 10-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF258
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

	LENGTH	1531 base pairs	
	TYPE:	nucleic acid	
	STRANDEDNESS:	single	
	TOPOLOGY:	linear	
	MOLECULE TYPE:	DNA (genomic)	
	FEATURE:		
	NAME/KEY:	CDS	
	LOCATION:	125..1453	
	US-08-833-963C-1		
Query Match	29.3%	Score 371.2;	DB 2; Length 1531;
Best Local Similarity	61.2%	Pred. No. 2.3e-106;	
Matches 617;	Conservative 0;	Mismatches 388;	Indels 3; Gaps 1;
Oy	262	TGCGCGTTGGATACCAAGATGGAAGAAAGCAACCAATGTTGGATGATGAGACAGTGGCA	321
Db	449	TGCGCCACAGGCTATGAGCCCGGACGATCAGGACAGCTGTGTGGATGTGGACAGTGTGCC	508
Oy	322	ACAGATTCCACACAGTGCACCCACCACCCAGATCTGCATCAATPACTGAAAGCGGGTACACC	381
Db	509	CAGGCCCTGCACGACTGTGGCCCGACGACGACTGCAATPACTTGGCTGGCTCCATACAG	568
Oy	382	TGCTTCCTGCACCGACGATATGGCTTGTGGAAGCCAGTCTTAGACATGATGAATGT	441
Db	569	TGCACCTGCTCCGTGATGTTACCCGACAGATCGGGCCGAGTGTGTGCATGTAGACAGTGC	628
Oy	442	CGATATGGTTTCTGCACGACAGCTCTGTGTCGAATGTCCGTGATCTTATCTTGTACATTC	501
Db	629	CGCTACGCGTACTGTCCACGACCGCTGTGGTGAACCTGCTGGCTCTTCGCTGCCAGTGC	688
Oy	502	AACCGTGGTTTAACTTCATGAGAGTGAAGAGTCTTGGCAAGATGTGAACAGTGTGCC	561
Db	689	GAGCCGGGCTTCCAGCTGGGGCTTACAAACCGCTCTGTGTGATGTAAGATGTGTAC	748
Oy	562	ACCGAAGAACCCCTGCGTGCACCACTGGTGTCAACACTTACGGCTCTTTCATCTGGCGGT	621
Db	749	ATGGGGGCCCATCGACGACCGCTGCTTCAACTCTATGAGGACCTTCTGTGTGGCTGC	808
Oy	622	GACCCAGATATGAATCTTGAGAGAGATGGCTTTCATTTGACGTATATGAGAGTGCAGC	681
Db	809	CACGAGGCGTATGAGCTGCATCGGGAGATGGCTTCTCTGCAGTGTATGATGATGTAGC	868
Oy	682	TTCCTGATGTTCTCTGCGCAACATGATAGTGTGAAGACCCCGGCACATATTCTGCTCC	741
Db	869	TACTCCAGCTACTCTGTCACTATACCCTGGCTTCAACGAGCCAGGCGCTTCTCTGCCAC	928
Oy	742	TGCCCTCAGGCTCATCTCTGCTGGATGACACGAAAGTCTCCAAAGATCAAGAAATGT	801
Db	929	TGCCACAGGTTTACAGCTGCTG--GCCACAGGCTCTGCCAAAGATTTGATGATGT	985
Oy	802	GAGCACAAGAACCAACAGTGCACCTGACGAGACGAGTGTCAAAATTTCAAAGGGGGCTTC	861
Db	986	GAGTCTGTGTGGCACACAGTGTCTCCGAGGCCCAAACTGTGTCAACTTCCATGAGGGGGCTAC	1045
Oy	862	AAATGCAATCGACCCCATCCGCTGTGAGGAGGCTTATCTGAGATCAGATGATACCGGCTGT	921
Db	1046	CGCTGCGTGAACACCAACCGCTGCTGGTGAAGCCCTTACATCAGAGTCTCTGAACACCGCTGT	1105
Oy	922	ATGTGTCTCTGTGAAGACCTTGGCTGTGACGAGACAGCCCTTACCATCTTGTACCGGGAC	981
Db	1106	CTGTGCGCGGCTCCAAACCTCTTATGTGACGAGACGCTTATCATTTGTGCACCGCTAC	1165
Oy	982	ATGGAACGTGTGTCAAGAGCGTCCGCTTCCGCTGACATCTTCCAAATTCAGCCACGACC	1041
Db	1166	ATGACCATCATCTTGAGAGCGGAGGTGCGCTGACGTTTCCAGTCCAGATCAGGCAACCTTC	1225
Oy	1042	CGCTACCTTGCGGGCTTTTACATTTTCCAGATTCAAATCTGGGAATGAGGGCAGAAATTT	1101
Db	1226	GTCATACCCGGTGTCTCAATGCTTTTCAATTCGCTGTGGAAATCTGCAGAGGGAGACTTT	1285
Oy	1102	TACATGGGCAACAGGGGCCCATATAGTGCACCTGTGGTATGACAGGCCCATCAAAGG	1161
Db	1286	TACATTTAGGCAAAATCAACAGTGTAGGCGCATGTGTGCTTCGCGCGGCGGTATACGGGC	1345

QY 61 TGCCGAACCATCCCGAGGCGCTGCGAGAGACATGATGTGTTAACCAAAATGGCGG 120
 |||||
 Db 517 TGCCGAACCATCCCGAGGCGCTGCGAGAGACATGATGTGTTAACCAAAATGGCGG 576
 QY 121 TATTTATGCAATCCCGAGCAAAACCTGTGTATCGAGGGCCCTACTCGAACCCCTACTCG 180
 |||||
 Db 577 TATTTATGCAATCCCGAGCAAAACCTGTGTATCGAGGGCCCTACTCGAACCCCTACTCG 636
 QY 181 ACCCCCTACTGAGTCCGTACCCAGAGCTGCCACCACCTCTCAGCTCCAAATCTATCC 240
 |||||
 Db 637 ACCCCCTACTGAGTCCGTACCCAGAGCTGCCACCACCTCTCAGCTCCAAATCTATCC 696
 QY 241 ACGATCTCCAGGCGCTTATATGCGCGCTTGGATACGATGGATGAAGCAACCAATGT 300
 |||||
 Db 697 ACGATCTCCAGGCGCTTATATGCGCGCTTGGATACGATGGATGAAGCAACCAATGT 756
 QY 301 GTGGATGTGACAGAGTGTGCAACAGATTCCACCAAGTGCACACCCACAGATCTGCATC 360
 |||||
 Db 757 GTGGATGTGACAGAGTGTGCAACAGATTCCACCAAGTGCACACCCACAGATCTGCATC 816
 QY 361 AATACGAAGGCGGTACACCTGCTCTGCAACGAGATTTGGCTTCTGGAAGGCCAG 420
 |||||
 Db 817 AATACGAAGGCGGTACACCTGCTCTGCAACGAGATTTGGCTTCTGGAAGGCCAG 876
 QY 421 TCGTTAGACATGTGATGTGCTGATGTTACTGACAGAGCTCTGTGCGAATTTCT 480
 |||||
 Db 877 TCGTTAGACATGTGATGTGCTGATGTTACTGACAGAGCTCTGTGCGAATTTCT 936
 QY 481 GGAATCTATTTCTGTACATGCACACCTGTTTACCTCAATGAGATGGAAGTCTTGC 540
 |||||
 Db 937 GGAATCTATTTCTGTACATGCACACCTGTTTACCTCAATGAGATGGAAGTCTTGC 996
 QY 541 CAAGATGTGACAGAGTGTGACACCGAGAACCCCTGCGTGAACCTGCTGTAACACCTAC 600
 |||||
 Db 997 CAAGATGTGACAGAGTGTGACACCGAGAACCCCTGCGTGAACCTGCTGTAACACCTAC 1056
 QY 601 GCGCTTTTCATCTGCGGCTGTGACCCAGATGATGAACTTGAAGATGGCTTCATTCG 660
 |||||
 Db 1057 GCGCTTTTCATCTGCGGCTGTGACCCAGATGATGAACTTGAAGATGGCTTCATTCG 1116
 QY 661 AATGATGTGACAGAGTGTGACCTTCTCTGAGTTCCTCTGCCAACATGAGTGTGAAACG 720
 |||||
 Db 1117 AATGATGTGACAGAGTGTGACCTTCTCTGAGTTCCTCTGCCAACATGAGTGTGAAACG 1176
 QY 721 CCCGACATCTTCTGCTCTGCTCCGCTCAGGGCTACATCGCTGATGAAACCAAGC 780
 |||||
 Db 1177 CCCGACATCTTCTGCTCTGCTCCGCTCAGGGCTACATCGCTGATGAAACCAAGC 1236
 QY 781 TGCCGAAGACATCAACGAATGTGAGACAGAGAACACACCTGCAACCTGACAGACGTGC 840
 |||||
 Db 1237 TGCCGAAGACATCAACGAATGTGAGACAGAGAACACACCTGCAACCTGACAGACGTGC 1296
 QY 841 TACCAATTTCAAGGGGGCTTCAAAATGATGACCCCATCGCTGTGAGAGCTTATCTG 900
 |||||
 Db 1297 TACCAATTTCAAGGGGGCTTCAAAATGATGACCCCATCGCTGTGAGAGCTTATCTG 1356
 QY 901 AAGATCAGTATTAACCGCTGTATGTCTGCTGAGAACCTGGCTGACAGAGACCAAGCC 960
 |||||
 Db 1357 AAGATCAGTATTAACCGCTGTATGTCTGCTGAGAACCTGGCTGACAGAGACCAAGCC 1416
 QY 961 TTTACCAATTTTACCGGGACATGACAGTGTGTGACAGAGCTCCGTTCCCGCTGATC 1020
 |||||
 Db 1417 TTTACCAATTTTACCGGGACATGACAGTGTGTGACAGAGCTCCGTTCCCGCTGATC 1476
 QY 1021 TTTCCAAATGCAAGCCAGACCGGCTACCTGGGGGCTATTACATTTTCCAGATCAAAATCT 1080
 |||||
 Db 1477 TTTCCAAATGCAAGCCAGACCGGCTACCTGGGGGCTATTACATTTTCCAGATCAAAATCT 1536
 QY 1081 GGGATGAGAGGAGAAATTTTACATGCGGCAAAAGGGGCCCATCAGTGTGCAACCTGCTG 1140
 |||||
 Db 1537 GGGATGAGAGGAGAAATTTTACATGCGGCAAAAGGGGCCCATCAGTGTGCAACCTGCTG 1596

QY 1141 ATGACAGCCCCCATCAAAAGGGCCCCGGAAATCCAGCTGACCTTGGAAATGATCATCTGC 1200
 |||||
 Db 1597 ATGACAGCCCCCATCAAAAGGGCCCCGGAAATCCAGCTGACCTTGGAAATGATCATCTGC 1656
 QY 1201 AACACTGTCACTCACTTCAAGAGGAGCTCCGTGTATCCGACGTGGGATATATGTGCGAG 1260
 |||||
 Db 1657 AACACTGTCACTCACTTCAAGAGGAGCTCCGTGTATCCGACGTGGGATATATGTGCGAG 1716
 QY 1261 TACCATTTC 1269
 |||||
 Db 1717 TACCATTTC 1725
 |||||
 RESULT 3
 US-09-188-930-67
 ; Sequence 67, Application US/09188930A
 ; Patent No. 6150502
 ; GENERAL INFORMATION:
 ; APPLICANT: Watson, James D.
 ; APPLICANT: Strachan, Lorna
 ; APPLICANT: Sleeman, Matthew
 ; APPLICANT: Murlson, James Greg
 ; APPLICANT: Onrust, Rene
 ; TITLE OF INVENTION: Compositions Isolated from Skin Cells
 ; FILE REFERENCE: 11000.1011c1
 ; CURRENT APPLICATION NUMBER: US/09/188,930A
 ; CURRENT FILING DATE: 1998-11-09
 ; NUMBER OF SEQ ID NOS: 348
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 67
 ; LENGTH: 1260
 ; TYPE: DNA
 ; ORGANISM: Rat
 ; US-09-188-930-67
 Query Match 66.3%; Score 841.8; DB 3; Length 1260;
 Best Local Similarity 89.4%; Pred. No. 5.1e-254;
 Matches 900; Conservative 5; Mismatches 102; Indels 0; Gaps 0;

QY 263 GCCGCTTTGGATACAGATGATGAAAGCAACCAATGTGTGATGTGAGAGTGTGCA 322
 |||||
 Db 7 GTCGCTTTGGATATGATGATGAAAGCAACCAAGTGTGTGATGTGAGAGTGTGCA 66
 QY 323 CAGATTTCCACCAAGTGTGCAACCCACCAATGTGTGATGATGAAAGGCGGTACACT 382
 |||||
 Db 67 CAGATTTCCACCAAGTGTGCAACCCACCAATGTGTGATGATGAAAGGCGGTACACT 126
 QY 383 GCTCTGCAACGAGGATTTGGCTTCTGGAAGGCGCAGTGTGATGATGATGATGATG 442
 |||||
 Db 127 GCTCTGCAACGAGGATTTGGCTTCTGGAAGGCGCAGTGTGATGATGATGATGATG 186
 QY 443 GCTATGTTACTGCAAGCAGCTGTGTGCAATGTTCTGATGATGATGATGATGATG 502
 |||||
 Db 187 GCTATGTTACTGCAAGCAGCTGTGTGCAATGTTCTGATGATGATGATGATGATG 246
 QY 503 AOCCTGTTTACCTCATGAGATGAGAGTCTTGTCCAAAGTGTGAAACAGAGTGTGCA 562
 |||||
 Db 247 AOCCTGTTTACCTCATGAGATGAGAGTCTTGTCCAAAGTGTGAAACAGAGTGTGCA 306
 QY 563 CCGAAGACCCCTGCTGCAAAACCTGCGTCAACACCAAGGCTTTCATCTGCGGCTGTG 622
 |||||
 Db 307 CTGAGAAACCCCTGTTTACAGACTGTGCTCAACACCTATGTGTTTTCATCTGCGGCTGTG 366
 QY 623 ACCGAGATTAAGACTTGAAGAGATGGGCTTCATTTGACAGTATGAGAGTGTGAGT 682
 |||||
 Db 367 ACCGAGATTAAGACTTGAAGAGATGGGCTTCATTTGACAGTATGAGAGTGTGAGT 426
 QY 683 TCTGTGATTTCTGTGCAACATGAGTGTGTGAACAGGCGGCGACATCTTGTGCTCT 742
 |||||
 Db 427 TCTGTGATTTCTGTGCAACATGAGTGTGTGAACAGGCGGCGACATCTTGTGCTCT 486
 QY 743 GCCCTCCAGGCTACATCTGTGTGATGAAACCGAAGGTGCCAAGATCAACGATATG 802

Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 CAGTGCAGAAATGGCTTTGACCTGGATGCGCCAGTGCAGTCTTTTATGATTTGTA 60
Db 457 CAGTGCAGAAATGGCTTTGACCTGGATGCGCCAGTGCAGTCTTTTATGATTTGTA 516
OY 61 TGGCGAACAATCCCGAGGCGTGGCGAGAGACATGATGTGTAAACCAAAATGCGGG 120
Db 517 TGGCGAACAATCCCGAGGCGTGGCGAGAGACATGATGTGTAAACCAAAATGCGGG 576
OY 121 TATTTATGCAATCCCGGACAAACCTTGTATGAGAGGCGCTTACTGCAACCCCTACTG 180
Db 577 TATTTATGCAATCCCGGACAAACCTTGTATGAGAGGCGCTTACTGCAACCCCTACTG 636
OY 181 ACCCCCTACTGAGGTCGTACCCAGACAGTCCGACACCTGCTCAAGCTCAACTTCC 240
Db 637 ACCCCCTACTGAGGTCGTACCCAGACAGTCCGACACCTGCTCAAGCTCAACTTCC 696
OY 241 ACATCTCCAGGCGCTTTATATGCGCTTTGGATACAGATGATGAAACCAACCAATGT 300
Db 697 ACATCTCCAGGCGCTTTATATGCGCGCTTTGGATACAGATGATGAAACCAACCAATGT 756
OY 301 GTGATGTGGACGAGTGTGCAACAGATTTCCACAGTGCACCCACCCAGATCTGCATC 360
Db 757 GTGATGTGGACGAGTGTGCAACAGATTTCCACAGTGCACCCACCCAGATCTGCATC 816
OY 361 AATACGAAAGGCGGTACACCTGCTCCTGACCCAGGATATGTGCTTGGAAAGGCCAG 420
Db 817 AATACGAAAGGCGGTACACCTGCTCCTGACCCAGGATATGTGCTTGGAAAGGCCAG 876
OY 421 TGCTTAGACATTTGATGATGTGCTATGCTTACTGCGACAGCTGTGCGAATGTCTT 480
Db 877 TGCTTAGACATTTGATGATGTGCTATGCTTACTGCGACAGCTGTGCGAATGTCTT 936
OY 481 GGATCTTATTTTGTATGATGCAACCCCTGTTTACCTTCATGAGATGGAAGTGTTCG 540
Db 937 GGATCTTATTTTGTATGATGCAACCCCTGTTTACCTTCATGAGATGGAAGTGTTCG 996
OY 541 CAAGATGTGAAGAGTGTGCGACAGAAACCCCTGCTGCAAAACCTGCTCAACACCTTC 600
Db 997 CAAGATGTGAAGAGTGTGCGACAGAAACCCCTGCTGCAAAACCTGCTCAACACCTTC 1056
OY 601 GGCTCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGAAAGTGGCTTCATTCG 660
Db 1057 GGCTCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGAAAGTGGCTTCATTCG 1116
OY 661 AGTGATATGAGAGTGTGAGCTTCTCTAGATTCCTTCCGACATGAGTGTGAACAG 720
Db 1117 AGTGATATGAGAGTGTGAGCTTCTCTAGATTCCTTCCGACATGAGTGTGAACAG 1176
OY 721 CCCGCGACATATCTGCTCTGCGCTCCAGGCTACATCTCTGATGACAACCGAAGC 780
Db 1177 CCCGCGACATATCTGCTCTGCGCTCCAGGCTACATCTCTGATGACAACCGAAGC 1236
OY 781 TGGCAAGACATCAAGAAATGTGAGACAGAAACACAGCTGCAACCTGAGACAGCTGC 840
Db 1237 TGGCAAGACATCAAGAAATGTGAGACAGAAACACAGCTGCAACCTGAGACAGCTGC 1296
OY 841 TACAATTTAAGAGGCGCTTCAATGATGATGCAACCCCATCCCTGTGAGAGGCTTATTCG 900
Db 1297 TACAATTTAAGAGGCGCTTCAATGATGATGCAACCCCATCCCTGTGAGAGGCTTATTCG 1356
OY 901 AGGATCAGTATACCGCTGTATGTCTGCTGAGAACCCCTGGCTGCAGAACCGGCC 960
Db 1357 AGGATCAGTATACCGCTGTATGTCTGCTGAGAACCCCTGGCTGCAGAACCGGCC 1416
OY 961 TTTTACATCTTTGACCGGACATGAGAGCTGTGTAGAGAGCTCCGCTTCGCTGACATC 1020
Db 1417 TTTTACATCTTTGACCGGACATGAGAGCTGTGTAGAGAGCTCCGCTTCGCTGACATC 1476
OY 1021 TTTCAATGCAAGCAGACCGGCTATGAGTGTGAGAGGCTTATTAATTTTCCAGATCAATCT 1080
Db 1477 TTTCAATGCAAGCAGACCGGCTATGAGTGTGAGAGGCTTATTAATTTTCCAGATCAATCT 1536

```

```

OY 1081 GGAATGAGGGCAGAGAAATTTTACATGCGGCAAAACGGGCGCCATCATGTCACCCCTG 1140
Db 1537 GGAATGAGGGCAGAGAAATTTTACATGCGGCAAAACGGGCGCCATCATGTCACCCCTG 1596
OY 1141 ATGACAGCCCAACAAAGGGCGCGGAAATCCAGCTGAGCTTGAAATGATCACTGTC 1200
Db 1597 ATGACAGCCCAACAAAGGGCGCGGAAATCCAGCTGAGCTTGAAATGATCACTGTC 1656
OY 1201 AACACTGTATCAACTTCAGAGGAGCTCCGTGATCCGACTCGGATATATGTGTGCGAG 1260
Db 1657 AACACTGTATCAACTTCAGAGGAGCTCCGTGATCCGACTCGGATATATGTGTGCGAG 1716
OY 1261 TACCATTTC 1269
Db 1717 TACCATTTC 1725

```

RESULT 2
US-09-212-168-2
; Sequence-2; Application US/09212168
; Patent No. 6303765
; GENERAL- INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Guebler, Karl J.
; TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/212,168
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/884,072
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0333 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2550 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: CORNMO101
; CLONE: 45517
; US-09-212-168-2

Query Match 100.0%; Score 1269; DB 4; Length 2550;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 CAGTGCAGAAATGGCTTTGACCTGGATGCGCCAGTGCAGTCTTTTATGATTTGTA 60
Db 457 CAGTGCAGAAATGGCTTTGACCTGGATGCGCCAGTGCAGTCTTTTATGATTTGTA 516

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 10:57:33 ; Search time 58.5574 Seconds

(without alignments)
6646.008 Million cell updates/sec

Title: US-09-674-379A-15

Sequence: 1 cagtcgcagcagtcgttga.....atgttcgcagtcaccattc 1269

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

IssuedPatents: 1
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/PCUS.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1269	100.0	2550	2	US-08-884-072-2
2	1269	100.0	2550	4	US-09-212-168-2
3	841.8	66.3	1260	4	US-09-188-930-67
4	371.2	29.3	1531	2	US-08-833-963C-1
5	371.2	29.3	2018	3	US-08-980-514-2
6	316.6	24.9	1018	3	US-09-188-930-259
7	241	19.0	2512	4	US-09-248-757-1
8	106.8	8.4	3373	2	US-08-897-443-2
9	89.4	7.0	401	4	US-09-643-597-278
10	80.2	6.3	220	4	US-09-404-879A-340
11	58.2	4.6	2461	1	US-08-282-141-1
12	53.2	4.2	3759	3	US-08-479-722B-3
13	52.4	4.1	5499	5	US-08-479-722B-1
14	52.4	4.1	5502	5	PCT-US95-02251-17
15	51.6	4.1	3753	5	PCT-US95-02251-2
16	51.6	4.1	4314	1	US-08-199-780-2
17	51.6	4.1	4314	2	US-08-316-650-2
18	51.6	4.1	5089	6	5177197-31
19	50.4	4.0	6412	4	US-08-652-877-17
20	50.4	4.0	6412	4	US-08-476-515A-17
21	50.4	4.0	14042	4	US-08-652-877-85
22	50.4	4.0	14042	4	US-08-652-877-89
23	50.4	4.0	14080	4	US-08-652-877-87
24	50.4	4.0	14083	4	US-08-476-515A-83
25	45.6	4.0	14086	4	US-08-652-877-83
26	45.6	3.6	3460	4	US-08-751-305-1
27	44.8	3.5	2181	6	5208144-36

28	43.4	3.4	3546	4	US-08-872-757-3	Sequence 3, Appl1
29	41.4	3.3	2452	4	US-09-381-779-1	Sequence 1, Appl1
30	39.8	3.1	3156	4	US-09-284-819-8	Sequence 8, Appl1
31	38	3.0	6048	4	US-09-634-920-3	Sequence 3, Appl1
32	37.8	3.0	1611	4	US-09-249-697A-2	Sequence 2, Appl1
33	37.8	3.0	1611	4	US-09-363-316B-2	Sequence 2, Appl1
34	37.8	3.0	2365	4	US-09-249-697A-5	Sequence 5, Appl1
35	37.8	3.0	2365	4	US-09-249-697A-18	Sequence 18, Appl1
36	37.8	3.0	2365	4	US-09-363-316B-5	Sequence 5, Appl1
37	37.8	3.0	2365	4	US-09-363-316B-23	Sequence 23, Appl1
38	37.2	2.9	2397	6	5258288-2	Patent No. 5258288
39	37	2.9	6344	4	US-08-843-417-1	Sequence 1, Appl1
40	35.4	2.8	825	1	US-08-312-870-6	Sequence 6, Appl1
41	35.4	2.8	1338	1	US-08-307-444A-8	Sequence 8, Appl1
42	35.4	2.8	1338	1	US-08-307-444A-8	Sequence 8, Appl1
43	35.4	2.8	1368	1	US-08-307-444A-7	Sequence 7, Appl1
44	35.4	2.8	1368	1	US-08-307-444A-7	Sequence 7, Appl1
45	35.4	2.8	1425	1	US-08-307-444A-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
Sequence 2 Application US/08884072
Patent No. 5208144
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/884,072
FILING DATE: Herewith
CLASSIFICATION: A24
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0333 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2550 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: CORNOTO1
CLONE: 45517
US-08-884-072-2
Query Match 100.0%; Score 1269; DB 2; Length 2550;
Best Local Similarity 100.0%; Pred. No. 0;